伤害厌恶的二因素结构及其抑制攻击的机制

岑雨珊[1] 夏凌翔[1] 黄润玉[2,1] 吕洁[1]

([1]西南大学心理学与社会发展研究中心，重庆 400715)([2]澳门大学社会科学学院心理学系，澳门 519000)

摘要 伤害厌恶是人类道德的核心，揭示它的结构与功能具有重要意义。伤害厌恶由伤害行为厌恶和伤害结果厌恶这两个相对独立的因素构成的观点尚未获得有力的实证支持，这两个因素在抑制攻击方面的差异及其背后的机制也不清楚。为了探究这些问题，我们实施了 1 个预实验和 3 个正式实验。预实验基于加工分离程序开发了伤害行为/结果厌恶分离任务，实验 1 使用该任务检验了伤害厌恶的结构，实验 2(包括 2a 和 2b)探究了伤害行为/结果厌恶抑制攻击的道德认知机制。研究结果显示，伤害厌恶由伤害行为厌恶和伤害结果厌恶这两个相对独立的因素构成。伤害结果厌恶可以通过道德推脱抑制攻击，但伤害行为厌恶对道德推脱和攻击的抑制作用不稳健。通过上述实验，本研究开发了伤害厌恶的研究工具，检验了伤害厌恶的二因素结构及其在抑制攻击认知与行为上的差异，揭示了伤害厌恶抑制攻击的道德认知路径。

关键词 伤害行为厌恶，伤害结果厌恶，加工分离程序，攻击，道德推脱

分类号 B849: C91

1 引言

伤害厌恶(harm aversion)指个体经历(实施、想到、看到或听到)伤害行为或结果时产生的不安反应或体验到不安的倾向(Cushman et al., 2013; Hou et al., 2023; Miller et al., 2014)。其是人类道德的核心(Yu et al., 2019)，具有减少不道德行为和增加亲社会行为的积极功能(Sarlo et al., 2014; Yu et al., 2021)。缺乏伤害厌恶则会增加个体的反社会行为(Crockett et al., 2014; Yu et al., 2021)。因此，探究伤害厌恶的本质与机制具有重要的科学价值与现实意义。学者们认为，伤害厌恶由伤害行为厌恶(harm action aversion)和伤害结果厌恶(harm outcome aversion)构成，而且这两个因素相互独立，具有不同的功能(Ariso, 2022; Cushman et al., 2012; Miller et al., 2014; Reynolds & Conway, 2018)。但也有研

收稿日期：2024-07-24

国家自然科学基金面上项目（32171071），重庆市自然科学基金面上项目（cstc2021jcyj-msxmX0607）和西南大学创新研究 2035 先导计划(SWUPilotPlan006)资助

通讯作者：夏凌翔, E-mail: xialx@swu.edu.cn

究显示，这两个因素的问卷分数的相关系数为 0.54~0.66(Miller et al., 2014)，并不支持它们相互独立的观点。此外，目前尚无实证研究探讨它们在抑制攻击的效应与机制上的差异。因此，伤害厌恶是否由两个相对独立的因素构成，以及这两个因素在抑制攻击上的异同仍不清楚，本研究将围绕这些问题展开研究。

1.1 伤害厌恶的结构

如前所述，伤害厌恶不是一个性质单一的概念，而是由伤害行为厌恶与伤害结果厌恶这两种在性质与功能上存在明显差异的心理因素构成(Cushman et al., 2012; Miller et al., 2014; Reynolds & Conway, 2018)。伤害行为厌恶指在不涉及伤害导致的消极后果的情况下，个体在经历(包括实施、想象、听到、看到等)伤害行为时产生的不安反应或体验到不安的倾向。伤害结果厌恶指在没有经历伤害行为的情况下，个体在经历(包括想象、听到、看到等)伤害导致的结果时产生的不安反应或体验到不安的倾向。

从理论上来说，伤害行为/结果厌恶的性质与根源存在明显差异。首先，根据道德二元论(theory of dyadic morality; Gray et al., 2012)，伤害行为厌恶是个体从施害者角度对伤害行为的感受与体验；伤害结果厌恶则是从受害者角度出发，涉及对受害者遭受伤害的共情。其次，根据 Blair(1995)提出的暴力抑制机制(violence inhibition mechanism, VIM)，Miller 等人(2014)和 Patil(2015)提出，伤害结果厌恶是个体与生俱来的，而伤害行为厌恶则很可能是后天习得的。总之，伤害行为/结果厌恶是两种在性质上存在明显差异的心理现象。

学者们进一步指出，伤害行为厌恶和伤害结果厌恶相对独立，不能构成一个高阶的伤害厌恶因素(Ariso, 2022; Miller et al., 2014)。然而，这一观点尚未获得有力的实证证据的支持。例如，基于上述观点编制的问卷——伤害行为/结果厌恶问卷(Miller et al., 2014)的调查结果显示，问卷两个维度的得分有较高的正相关(0.54~0.66)(Miller et al., 2014; Reynolds & Conway, 2018)。该结果与这两个因素相对独立、不能形成一个高阶因素(伤害厌恶)的观点矛盾。对此，Miller 等人(2014)的解释是，被试在完成伤害行为厌恶分量表的题项时，可能会不自觉地想到伤害结果，因此该分量表的内容也涉及伤害结果厌恶。可见，有效分离伤害行为与结果厌恶是检验上述观点的一个关键环节。加工分离程序(process dissociation procedure, PDP)可能提供了一种解决方案。使用该范式，我们可以通过两种指标来检验伤害行为/结果厌恶的相对独立性：(1)二者不相关或弱相关(Armstrong et al., 2018; Conway et al., 2024; Conway & Gawronski, 2013)；(2)二者在抑制攻击的功能上存在明显差异(Carroll et al.,

2018; Crick & Dodge, 1996; Fung, 2021; Vaughan et al., 2023)。

1. 2 伤害行为/结果厌恶的分离

Jacoby (1991)最早使用 PDP 分离有意识注意和无意识注意。之后的学者采用该范式计算出功利主义参数(U)和道义主义参数(D)，并成功实现了二者的分离(Armstrong et al., 2018; Conway et al., 2024; Conway & Gawronski, 2013)。

如果伤害行为厌恶与伤害结果厌恶也是两种相对独立的道德心理，那么 PDP 就应该可以分离它们。PDP 的基本操作是形成不同心理成分导致不同反应的不相容条件(incongruent condition)和导致相同反应的相容条件(congruent condition)。通过这两种条件下的任务，我们可以量化每个心理成分对相关反应的相对贡献，并计算出能代表这两种心理成分的参数，从而实现对这两种心理成分的分离。根据伤害行为/结果厌恶的界定和性质，它们对两类伤害相关的反应(避免直接实施伤害行为与避免直接观看伤害结果)具有相对独立的作用(Ariso, 2022; Cushman et al., 2012; Miller et al., 2014)，符合使用 PDP 的基本条件。具体来说，在现实生活中存在可以让高伤害行为厌恶者和高伤害结果厌恶者做出相同选择的相容条件和做出不同选择的不相容条件。

个体在相容与不相容条件下的选择可以借助图 1 中的处理树说明：(1) 伤害结果厌恶(O 参数)主导的选择(顶部路径)；(2) 伤害行为厌恶(A 参数)主导的选择(中间路径)；(3) 伤害行为厌恶(A 参数)和伤害结果厌恶(O 参数)都不主导的选择(底部路径)。此外，“1–O”表示伤害结果厌恶(O 参数)不主导选择的情况，“1–A”表示伤害行为厌恶(A 参数)不主导选择的情况。当伤害结果厌恶(O 参数)主导选择时，在相容条件下，个体不会选择伤害行为，但在不相容条件下，就会选择伤害行为。当伤害结果厌恶(O 参数)不主导选择，而是伤害行为厌恶(A 参数)主导选择时，在相容条件和不相容条件下，个体都不会选择伤害行为。在伤害结果厌恶(O 参数)与伤害行为厌恶(A 参数)都不主导选择的情况下，在相容和不相容条件下，个体都更可能选择做出伤害行为。

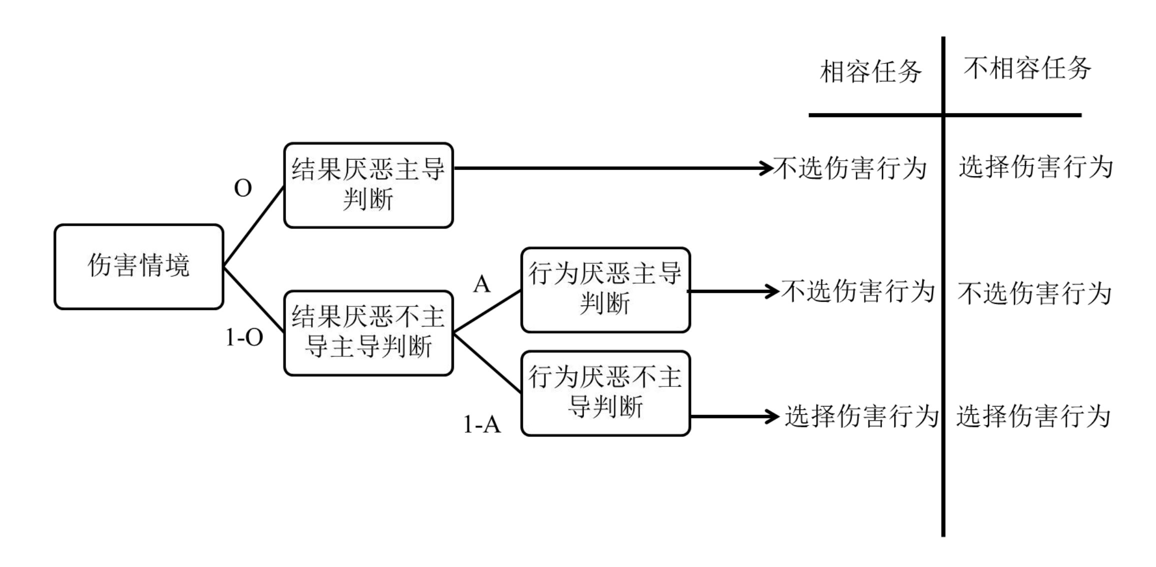


图 1 伤害行为厌恶与伤害结果厌恶分离的处理树

(注：A，伤害行为厌恶参数；O，伤害结果厌恶参数)

因此，我们可以用伤害行为厌恶(A 参数)和伤害结果厌恶(O 参数)表征相容和不相容条件下选择伤害行为和不选伤害行为的概率(p)：

① p (不选伤害行为|相容条件) = O + [(1–O)×A]

② p (选伤害行为|相容条件) = (1–O) × (1–A)

③ p (不选伤害行为|不相容条件) = (1–O) × A

④ p (选伤害行为|不相容条件) = O + [(1–O) × (1–A)]

基于这四种情况下个体选择伤害行为和不选伤害行为的概率(p)，我们可以计算出 A 参数和 O 参数。

⑤O = p (不选伤害行为|相容条件) – p (不选伤害行为|不相容条件)

⑥A = p (不选伤害行为|不相容条件) / (1–O)

借助不相容和相容条件下的选择任务，我们可以获得分别表征伤害行为厌恶和伤害结果厌恶的参数。根据已有研究的观点(Conway et al., 2024; Conway & Gawronski, 2013)，这两个参数不相关或弱相关就表明伤害行为厌恶和伤害结果厌恶是两个相对独立的变量。

1.3 伤害行为/结果厌恶抑制攻击的功能与机制

1.3.1 伤害行为/结果厌恶对攻击的抑制作用

尽管不少学者(Crockett et al., 2010; Crockett et al., 2015; Martinez et al., 2024; Perera et al., 2016)

指出伤害厌恶可以抑制攻击，但是仍缺乏充分的实证证据支持这一观点，并且尚未有研究揭示伤害厌恶抑制攻击的机制。因此，本研究将探究伤害行为/结果厌恶对攻击的抑制作用及其背后的中介机制。

攻击是一种存在伤人意图的伤害反应或倾向，而且这种伤害是对方想要回避的(李芮, 夏凌翔, 2021)。伤害行为厌恶是一种对伤害行为的反感和抵触情绪，它会导致个体反感攻击这种故意伤害他人的行为。因此，伤害行为厌恶很可能会抑制个体的攻击行为(Martinez et al., 2024; Miller et al., 2014; Miller & Cushman, 2013)。伤害结果厌恶可以抑制攻击是因为它会引发个体对实施攻击行为的回避反应。有一些间接证据(Blair, 1995; Buss, 1966; Vollberg & Cikara, 2024) 支持该观点。由于伤害行为厌恶和伤害结果厌恶的功能存在差异(Miller et al., 2014; Reynolds & Conway, 2018; Zapata & Deroy, 2023)，我们可以推论，这两个因素对攻击的抑制作用很可能存在明显差异。

1.3.2 伤害行为/结果厌恶抑制攻击的道德认知路径

伤害行为/结果厌恶是重要的道德因素(Crockett et al., 2015; Yu et al., 2019)，因此，道德路径很可能是它们抑制攻击的重要机制。在影响攻击的道德认知因素中，道德推脱(moral disengagement)是目前最受认可的一个变量(Cen et al., 2022; Ogunfowora et al., 2022)。道德推脱指个体为了减少或避免道德抑制和自我谴责，将不道德的行为认知重构为可接受的或道德的反应或倾向(Bandura, 2002; Li & Xia, 2024)。一些研究(Chowdhury & Fernando, 2013; Leviston & Walker, 2020)提示，伤害行为/结果厌恶与道德推脱的关系密切。因此，道德推脱很可能是伤害行为/结果厌恶抑制攻击的一种道德认知路径。

伤害行为/结果厌恶能抑制道德推脱，主要是因为道德情绪可以抑制包括道德推脱在内的消极道德认知。例如，共情、内疚被认为可以抑制道德推脱(Chowdhury & Fernando, 2013; Leviston & Walker, 2020; Shen et al., 2023)。具体来说，伤害行为/结果厌恶是对伤害行为、意图或结果的反感(Ariso, 2022; Cushman et al., 2012; Miller et al., 2014)，因此很可能会抑制个体从积极角度对伤害行为、意图或结果进行认知扭曲或美化。从积极角度扭曲或美化伤害行为与意图则是攻击相关的道德推脱的典型表现(Bjärehed et al., 2019; Falla et al., 2023)。据此我们可以推论，伤害厌恶很可能会抑制个体进行攻击相关的道德推脱。

道德推脱会使道德体系的调节作用失效，削弱个体的道德自我调节功能(Wang et al., 2022)，从而减少个体的道德系统对攻击的抑制作用，甚至在道德上认可伤害行为(Bjärehed et al., 2019;

Thornberg et al., 2022)。因此，道德推脱可以起到驱动攻击的作用(Guo et al., 2024)。

在此基础上，我们进一步推论，伤害行为/结果厌恶会抑制个体从道德上对攻击行为进行辩解或美化，并进而减少个体实施攻击的可能性或频率。

1.4 本研究

综上所述，本研究提出三个假设。假设 1：伤害厌恶由伤害行为厌恶与伤害结果厌恶这两个相对独立的变量构成，这两个变量可以用 PDP 分离，并分别用 A 参数和 O 参数表示。假设 2：伤害行为/结果厌恶都对攻击具有抑制作用，但二者的抑制功能存在差异。假设 3：道德推脱在伤害行为/结果厌恶和攻击的关系之间起到中介作用。

为了检验上述假设，本研究首先基于 PDP 开发伤害行为/结果厌恶分离任务，获得代表伤害行为厌恶与伤害结果厌恶的参数。之后，我们通过效度检验来考察这两个参数是否能够代表伤害行为/结果厌恶，同时考察这两个参数之间的相关。最后，我们基于这两个参数进一步探究伤害行为/伤害结果厌恶对攻击的抑制作用，以及道德推脱在其中的中介作用。

为了使本研究的统计分析更为可靠，根据已有文献的建议(Wagenmakers et al., 2011)，对于假设 1 涉及到的无差假设(H_0)，本研究不再使用传统的相关分析，而是采用贝叶斯相关分析，并报告贝叶斯因子(Bayes factor, BF)。当 $BF_{10} > 10$、处于 3~10 以及 1~3 时，分别表明有较强、中等、较弱的证据支持备择假设(H_1)；当 $BF_{10} < 0.10$、处于 0.10~0.33 以及 0.33~1 时，分别表明有较强、中等、较弱的证据支持无差假设 (Wagenmakers, Marsman et al., 2018)。

2 预实验：伤害行为/结果厌恶分离任务的开发

应用 PDP 分离伤害行为厌恶和伤害结果厌恶的关键是开发出具备相容条件和不相容条件的任务。我们借鉴道德两难任务(Armstrong et al., 2018; Conway et al., 2024; Conway & Gawronski, 2013)开发出基于 PDP 的伤害行为/结果厌恶分离任务。

此外，与伤害行为/结果厌恶问卷(Miller et al., 2014)相比，伤害行为/结果分离任务的另一个特点是，它主要针对的是暴力伤害厌恶。伤害行为/结果厌恶问卷中有许多题项针对意外伤害之类的非暴力伤害的厌恶体验。这导致该问卷的分数与攻击的相关很低(Hou et al., 2023)。伤害厌恶是一种道德情绪，其核心成分是对暴力伤害的厌恶。因此，我们在开发伤害行为/结果分离任务材料时，主要围绕暴力伤害进行。

2.1 方法

2.1.1 被试

通过在教室发放纸质问卷的方式，我们在线下招募 20 名本校学生，同时通过问卷星平台在线上招募 22 名在校大学生，共 42 人。所有被试在完成伤害行为/结果分离任务中的反应均未违反 PDP 计算规则。排除 2 名作答不认真的被试(作答时间小于 3 分钟)，本实验最终获得 40 个有效被试(男生 19 人，女生 21 人)，年龄为 18.08~26.83 岁(M = 21.52, SD = 2.68)。实验结束后，每位被试获得 6 元报酬。基于敏感性分析，本实验设置统计检验力为 0.80，进行双尾检验(α = 0.05)，结果显示，40 人的样本量检验到的最小效应为 r = 0.41。为了检验线上数据和线下数据的同质性，我们进行了独立样本 t 检验。结果显示，A 参数(t = 0.76, p = 0.453, Cohen's d = 0.25, BF_{10} = 0.39)、O 参数(t = -0.58, p = 0.565, Cohen's d = -0.19, BF_{10} = 0.36)、相容情境下的选择难度(t = 0.69, p = 0.497, Cohen's d = 0.22, BF_{10} = 0.38)、不相容情境下的选择难度(t = 0.20, p = 0.840, Cohen's d = 0.06, BF_{10} = 0.32)和想象难度(t = 0.32, p = 0.751, Cohen's d = 0.10, BF_{10} = 0.32)在这两种施测方式上均无显著差异。

2.1.2 研究材料和程序

伤害行为/结果厌恶分离任务主要由情境材料构成。每个情境材料包括三个部分：故事情境、不相容条件和相容条件。不相容条件和相容条件各有两个选项。我们采用以下步骤完成开发：第一，基于伤害行为/结果厌恶问卷(Miller et al., 2014)和道德两难情境(Armstrong et al., 2018; Conway & Gawronski, 2013)收集“仅表现伤害行为但不体现伤害结果” 和“仅体现伤害结果但不表现伤害行为”的典型例子与故事情境。第二，根据对 5 名非心理学专业大学生的访谈补充编写故事情境，形成了 32 个类型不同的故事情境。第三，邀请 6 名心理学专业的大学生完成这 32 个情境任务，评定情境想象和做出行为选择的难易程度，并提出修改建议。第四，根据评定结果和修改建议，我们讨论、筛选和修改情境材料。例如，我们对有偏向性的故事情境及其选项进行了修改。第五，进行 3 次小样本(11~16 人)施测，根据施测结果不断修改和调整实验材料。最终，我们确定了 10 个情境材料(详见附录 1)。

完成伤害行为/结果厌恶分离任务，被试需要在 10 个故事情境的不相容和相容条件下分别进行选择。故事情境与具体条件采用伪随机的方式呈现(即相容条件或不相容条件不能连续出现超过两次，且同一个故事情境下的相容和不相容条件不能接连出现)，选项的顺序在被试间进行平衡。之后，被试需要采用李克特 7 点量表评定进行情境想象和做出行为选择的难易程度，并给出对实验材料的

修改建议。

2.2 结果与讨论

实验结果表明，被试在相容条件下选择伤害行为的比例为 38.75%，在不相容条件下选择伤害行为的比例 57.50%。这一结果与 Conway 和 Gawronski (2013)采用 PDP 分离道义主义和功利主义时获得的结果类似。

根据引言中的公式 5 和 6，我们计算出了代表伤害行为厌恶和伤害结果厌恶的 A 参数与 O 参数。进一步的贝叶斯相关分析显示，有较弱的证据支持这两个参数的相关不显著($BF_{10} = 0.46$)。这与已有研究的结果(Conway et al., 2024; Conway & Gawronski, 2013)相符。

此外，有较弱的证据支持被试评定的相容条件($Min = 1$, $Max = 7$, $M = 3.48$, $SD = 1.15$)与不相容条件($Min = 1$, $Max = 7$, $M = 3.27$, $SD = 1.16$)下的选择难度不存在显著差异($t(39) = 1.80$, $p = 0.080$, Cohen's $d = 0.28$, $BF_{10} = 0.74$)。被试对情境故事的想象难度的评分均值在“比较简单”至“中等”之间($Min = 1$, $Max = 7$, $M = 3.78$, $SD = 1.80$)。这提示，不相容条件下的选择难度和相容条件下的选择难度基本相当，而且被试均能较好的想象和代入故事情境。

综上所述，预实验的结果表明，基于 PDP 开发的伤害行为/结果厌恶分离任务可以获得相对独立的伤害行为厌恶与伤害结果厌恶的参数，初步支持了假设 1。

根据被试的完成与评定情况以及他们提出的相关建议，我们再次讨论了每个情境材料的表述，在不改变意思的前提下对部分语言表达进行了优化。最终形成了正式的伤害行为/结果厌恶分离任务。

预实验的样本量较小，因此，我们将在实验 1 中使用较大的样本量来进一步检验 PDP 分离伤害行为/结果厌恶的有效性。

3 实验 1 伤害厌恶的二因素结构

实验 1 拟使用伤害行为/结果厌恶分离任务进一步检验伤害行为厌恶与伤害结果厌恶是两个相对独立的因素的假设。首先，我们将考察代表伤害行为/结果厌恶的 A 参数和 O 参数的相关。其次，根据已有研究的观点和结果(Almeida et al., 2024; Contreras-Huerta et al., 2022; Cushman et al., 2013; Li, 2023; Martinez et al., 2024; Miller et al., 2014; Su & Xia, 2024)，我们使用伤害行为/结果厌恶问卷测量的伤害行为厌恶与伤害结果厌恶，以及共情、道德判断、道义主义 D 参数、精神病态、身体攻

击、口头攻击等为效标变量来检验 A 参数和 O 参数的效度。

3.1 被试

我们采用 G*Power 估计样本量，根据学者们对统计检验力(吴艳，温忠麟, 2011a)和社会心理效应量(Richard et al., 2003)的建议，将统计检验力设为 0.80，将效应量设置为 r = 0.21，进行双尾检验(α = 0.05)，结果显示，所需的样本量为 173 人。因此，本实验在三所高校共招募了 347 名大学生。根据以下排除标准：(1)作答时间小于 5 分钟，(2)大面积(超过 2/3 的题项)规律性作答，(3) 4 道陷阱题(如“这道题请选择完全符合”)全部回答错误，我们排除了作答不认真的 6 名被试。此外，根据 PDP 的规则(Conway & Gawronski, 2013)，本实验删除了在相容条件中都选择不伤害的被试 54 人。最终获得有效被试 287 人(男生 103 名，女生 184 名)，年龄为 18.08~23.75 岁(M = 20.32, SD = 1.30)。实验结束后每位被试获得 6 元报酬。

实验 1 将统计检验力设置为 0.80 进行敏感性分析，结果显示，当前样本量能检验的最小效应量为 r = 0.16。

3.2 研究材料

3.2.1 伤害行为/结果厌恶分离任务

实验 1 采用预实验开发的伤害行为/结果厌恶分离任务来获得伤害行为厌恶与伤害结果厌恶的参数。该任务总共有 10 个情境故事，每个情境都包括不相容条件与相容条件，每种条件下设置两个选项。该 10 个情境故事在不相容条件与相容条件下都采用伪随机顺序呈现，呈现标准为：(1)不连续呈现同一故事情境的相容条件和不相容条件，(2)相容条件或不相容条件不能连续出现两次以上；同时，选项顺序在被试间进行平衡。

3.2.2 伤害行为/结果厌恶问卷

伤害行为/结果厌恶的问卷分数采用 Miller 等人(2014)编制的伤害行为/结果厌恶问卷(Harm Action/Outcome Questionnaire)获得。该量表共有 34 个题目，其中的 9 个题项测量伤害行为厌恶(例如，“在一个演出中，你用一把可伸缩的道具刀刺向一位演员的脖子”)，14 个题项测量伤害结果厌恶(例如，“你看见有人关轿车门时夹住了自己的手指”)，剩余的 11 个是控制题项。量表采用 7 点评分，被试用从 1(完全没有)到 7(非常强烈)评定自己体验到不安的程度，得分越高代表个体的伤害行为/结果厌恶的水平越高。该问卷的英文版(Miller et al., 2014; Patil, 2015)和中文版(Hou et al., 2023)均有良

好的信度和效度。在本研究中，该量表的结构效度良好: χ^2/df = 49.83/19, RMSEA [90% C.I.] = 0.07[0.05, 0.10], CFI = 0.98, TLI = 0.96, SRMR = 0.03，本研究中伤害行为/结果厌恶分量表的 Cronbach's α分别为 0.88 和 0.92。

3.2.3 共情分量表

共情的测量采用的是 Davis (1983)的人际反应指数量表(Interpersonal Reactivity Index , IRI)中的共情关注(empathic concern)分量表，该分量表包含 7 个题项("对那些不幸的人，我通常会感到心疼和担忧")。该量表请被试采用从 1(完全不符合)到 5(完全符合)的方式评定题项描述的内容与自己情况的符合程度。已有研究表明该量表的英文版(Davis, 1983)和中文版(戎幸 等, 2010)均具有良好的信度和效度。在本研究中，该分量表的 Cronbach's α为 0.67。

3.2.4 道德两难任务

实验 1 采用 Conway 和 Gawronski(2013)开发的道德两难任务来获得道义主义参数 D。该任务共有 10 个道德两难情境故事，每个情境都有两个版本分别对应不相容条件和相容条件。在不相容和相容条件下，该任务都会给出 1 种伤害行为选择，请被试判断选择该伤害行为是否可以接受。不相容条件下的是一种可以获得很大利益的伤害行为，而相容条件下的伤害行为获得利益相对很小。在不相容条件下，高道义主义者会倾向于不接受伤害行为，而功利主义者则倾向于接受该伤害行为。在相容条件下，高功利主义和高道义主义者都倾向于不接受伤害行为。

3.2.5 莱文森精神病态量表

莱文森精神病态量表(Levenson Self-Report Psychopathy Scale; Levenson et al., 1995)被用于测量个体的精神病态特质。该量表有 19 个题项，如"在今天这个世界，我觉得为了取得成功，做任何事都是正当的"。被试根据自己的实际想法采用 1(完全不符合)到 4(完全符合)的评分方法进行选择。已有研究显示，该量表的英文版(Levenson et al., 1995)和中文版(邓俏文 等, 2017)均具有良好的信度和效度。在本研究中，该量表的 Cronbach's α为 0.80。

3.2.6 身体攻击与口头攻击分量表

本研究采用 Buss-Perry 攻击问卷(Buss-Perry Aggression Questionnaire, BPAQ; Buss &Perry, 1992)中的身体攻击和言语攻击分量表来测量个体的攻击行为，其中身体攻击包含 9 个题项(例，"如果受

到多次挑衅，我就会打人")，言语攻击包含 5 个题项(例，"当有人来烦我的时候，我就容易骂人")。问卷采用 5 点计分，1 代表"完全不符合"，5 代表"完全符合"。该量表的英文版(Buss & Perry, 1992)和中文版(Quan et al., 2019)均具有良好的信度和效度。本研究中身体攻击和口头攻击分量表的 Cronbach's α分别为 0.82 和 0.72。

3.3 研究程序

请被试在线下完成伤害行为/结果厌恶分离任务、伤害行为/结果厌恶问卷、人际反应指数量表中的共情分量表、道德两难任务、莱文森精神病态量表、Buss-Perry 攻击问卷中的身体攻击和言语攻击分量表，以及人口统计学信息问卷。

3.4 统计处理

实验 1 借助 SPSS 24.0 对数据进行分析。首先，我们对相容条件和不相容条件下的选择比例进行了分析。相容与不相容条件下的选择比例的计算方法为被试选择实施伤害行为的次数除以相容/不相容情境的总个数(10 个)。其次，依据 PDP 程序的计算规则(见引言 1.2 部分的计算方式)，我们获得 A 参数和 O 参数，并借助 JASP 0.19.1(Wagenmakers, Love et al., 2018)对这两个参数进行贝叶斯相关分析。再次，我们采用 Pearson 相关计算 A 参数、O 参数与效标变量的相关，并计算了不显著结果的贝叶斯因子。最后，采用 R 包的 Silver's Z 程序(Diedenhofen & Musch, 2015)的 Pearson 和 Filon (1898) *z* 检验，我们对 A 参数、O 参数与效标变量的相关系数进行相关差异显著性检验(计算网址为：http://comparingcronbachalphas.org/)。

3.5 结果与讨论

3.5.1 伤害行为选择比例分析

在不相容与相容条件下，被试选择实施伤害行为的次数占总选择次数的比例分别为 62.40% (SD = 0.18) 和 27.98% (SD = 0.18)。这一比例与 Conway 和 Gawronski(2013)的研究中个体在相容与不相容条件下选择做出伤害行为的比例接近。进一步的配对样本 t 检验显示，个体在不相容条件做出伤害行为的选择显著高于相容条件，$t(286) = 20.80$, $p < 0.001$, Cohen's $d = 2.46$。

3.5.2 参数效度检验

表 1 A 参数、O 参数与效标之间的相关

	A 参数	O 参数	相关差异检验 z	相关差异检验 p
伤害行为厌恶效标	0.16^{**}	0.23^{***}	–0.90	0.370
伤害结果厌恶效标	0.08	0.23^{***}	$–2.02^{*}$	0.044
共情关注	0.13^{*}	0.30^{***}	$–2.25^{*}$	0.024
传统道德判断	0.27^{***}	0.22^{***}	0.61	0.544
道义主义 D 参数	0.29^{***}	0.28^{***}	0.22	0.830
功利主义 U 参数	–0.05	–0.01	–0.56	0.579
莱文森精神病态	$–0.27^{***}$	$–0.42^{***}$	2.07^{*}	0.038
身体攻击	$–0.22^{***}$	$–0.22^{***}$	0.05	0.958
口头攻击	–0.09	$–0.13^{*}$	0.47	0.638

注：A 参数指伤害行为厌恶参数，O 参数指伤害结果厌恶参数；* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$。

此外，A 参数和 O 参数相关的贝叶斯因子为 $BF_{10} = 0.27$，表明有中等程度的证据支持二者不相关，这提示 A 参数和 O 参数所代表的两种心理成分很可能是相对独立的，支持了假设 1 以及学者们的相关观点(Ariso, 2022; Cushman et al., 2012; Miller et al., 2014)。

A 参数(代表伤害行为厌恶)和 O 参数(代表伤害结果厌恶)与效标变量的相关分析的结果如表 1 所示。由于 A 参数与伤害结果厌恶的问卷得分、功利主义 U 参数和口头攻击的相关，以及 O 参数与功利主义 U 参数的相关均不显著，我们进行了贝叶斯相关分析，结果显示，有中等到较强程度的证据支持零假设。

本实验的结果支持了伤害厌恶由伤害行为厌恶与伤害结果厌恶这两个相对独立的变量构成的假设。第一，伤害行为厌恶的问卷分数与 A 参数显著正相关，伤害结果厌恶的问卷分数与 A 参数的相关不显著($r = 0.08$, $p = 0.201$, $BF_{10} = 0.17$)。此外，伤害行为厌恶的问卷分数与 O 参数显著正相关，这支持了伤害行为/结果厌恶问卷的伤害行为厌恶分量表涉及了伤害结果厌恶的观点(Miller et al., 2014)。第二，共情关注与 A 参数、O 参数均显著正相关；进一步的差异显著性检验显示，共情关注与 A 参数之间的相关系数显著小于与 O 参数之间的相关系数($z = –2.25$, $p = 0.024$)。这些结果与已有研究发现的共情与伤害结果厌恶关系更为密切的结果(Almeida et al., 2024; Cushman et al., 2012; Miller et al., 2014)一致。第三，精神病态与 A 参数、O 参数均呈显著负相关；进一步的差异显著性检验显示，精神病态与 O 参数的相关系数显著大于与 A 参数的相关系数($z = 2.07$, $p = 0.038$)。这些结果与已有研究发现的相对于伤害行为厌恶，伤害结果厌恶与精神病态的负向关系更强的结果(Patil,

2015; Reynolds & Conway, 2018)一致。第四，道德判断与 A 参数、O 参数均呈显著正相关。这些结果与已有研究发现的道德判断与伤害行为/结果厌恶均呈正相关的结果(Reynolds & Conway, 2018; Li, 2023; Wiech et al., 2013)一致。第五，道义主义 D 参数与 A 参数、O 参数均呈显著正相关。这些结果与已有研究发现的道义主义参数与伤害行为/结果厌恶均呈正相关的结果(Cushman, 2013; Reynolds & Conway; 2018)相符。此外，功利主义 U 参数与 A 参数(r = -0.05, p = 0.391, BF_{10} = 0.11)、O 参数(r = -0.01, p = 0.911, BF_{10} = 0.11)的相关不显著。第六，身体攻击与 A 参数、O 参数均呈显著负相关；口头攻击与 O 参数的负相关显著，但口头攻击与 O 参数的相关值小于目前样本可检验的最小效应量，因此我们需要进一步探讨 O 参数和口头攻击之间的关系。

此外，本研究获得了一些意外的结果。例如，A 参数、O 参数与对应的伤害行为/结果厌恶的问卷分数的相关较低(0.16~0.23)。如前所述，伤害行为/结果厌恶问卷(Miller et al., 2014)涉及了不少非暴力伤害厌恶成分，而伤害行为/结果厌恶分离任务主要针对的是暴力伤害为厌恶。这可能是出现该种结果的主要原因。此外，有中等程度的证据显示 A 参数与口头攻击的相关不显著(r = -0.09, p = 0.132, BF_{10} = 0.23)。这可能是因为口头攻击带来的伤害通常相对较低而且具有主观性(即受害者在意时才会带来伤害，不在意就基本不会导致伤害)，因此人们对口头攻击行为的伤害厌恶也相对较弱。这个结果提示，伤害行为厌恶对不同种类的攻击的抑制功能很可能存在差异。

4 实验 2 伤害行为/结果厌恶抵御攻击的道德认知机制

如前所述，虽然学者们普遍认为伤害厌恶可以抑制攻击，但伤害厌恶抵御攻击的实证证据并不充分，抵御机制也不清楚。因此，实验 2 尝试在实验 1 的基础上探索伤害行为/结果厌恶对攻击的抵御作用以及道德推脱在其中的中介作用。

4.1 实验 2a 道德推脱在伤害行为/结果厌恶与攻击关系中的中介作用

4.1.1 被试

由于实验 2a 的中介变量与因变量是采用问卷测量，此类变量的关系通常具有小至中等的效应量(Zhao & Seibert, 2006)，因此依据 Cohen (1969)提出的 f^2 值建议，实验 2a 设置效应量为f^2 = 0.07，并采用 G*Power 估计样本量，在 5 个预测因子的回归分析中(A 参数、O 参数、道德推脱、性别、年龄)，α = 0.05，计算得到 189 人能达到 0.80 的统计检验力。然而，本研究中的模型可能存在两个

因变量，因此需要 378 名被试。

本研究在四所高校共招募了 518 名大学生，排除作答时间小于 5 分钟以及作答不认真的 17 名被试，根据分离范式的方程删除在相容条件中都选择不伤害的被试 78 人，最终有效被试为 423 人(男生 170 名，女生 253 名)，年龄为 18.08~23.75 岁(M = 20.32, SD = 1.30)，调查结束后获得 6 元报酬。我们采用 G*Power 进行敏感性分析，结果显示，423 人的样本量能检测的最小效应量为 f^2 = 0.03。

4.1.2 研究程序

受到 COVID-19 的影响，本研究通过线上线下结合的方式请被试完成伤害行为/结果厌恶分离任务、公民道德推脱问卷以及 Buss-Perry 攻击问卷中的身体攻击和口头攻击分问卷。被试作答平均时间为 10 分钟。

4.1.3 测量工具

伤害行为/结果厌恶分离任务：同实验 1。

公民道德推脱问卷：实验 2a 采用 Caprara 等人(1996)编制的公民道德推脱问卷(Civic Moral Disengagement Questionnaire)，该问卷共有 32 个题目(如“为了保护自己的利益，通常必然会使用到武力”)。问卷采用五点计分，评分从 1 分(非常不同意)到 5 分(非常同意)。在国内外研究中(Caprara et al., 2009; 王兴超 等, 2013)，该量表的信度与效度良好。在本研究中，该量表 Cronbach’s α为 0.928。

身体攻击和言语攻击分量表：与实验 1 一致，在本研究中的口头攻击(Cronbach’s α = 0.75)和身体攻击(Cronbach’s α = 0.84)的信度良好。

4.1.4 统计处理

实验 2a 借助 SPSS 24.0 和 Mplus 8.0 对数据进行分析。首先，我们对变量进行描述统计，计算变量间的相关，并对 A 参数与 O 参数的关系以及相关不显著的变量关系采用 JASP 进行贝叶斯相关分析。其次，我们进行了 Little 完全随机缺失检验(MCAR; Little & Rubin, 2002)。结果表明缺失值模式为完全随机缺失(χ^2 = 1758.33, df = 1874, p = 0.972)，因此使用完全信息极大似然(full information maximum likelihood, FIML) 来处理缺失值。由于研究变量的数据未达到多重正态分布的标准，因此按照已有文献的建议(Gibson & Ninness, 2005)，使用稳健极大似然法(robust maximum likelihood)作为估计方法。再次，我们采用误差相关法对问卷题项进行打包(吴艳, 温忠麟, 2011b)，使每个研究变量

有四个项目包；最后，我们进行结构方程模型分析，并采用抽取 5000 次的偏差矫正的 Bootstrap 方法检验道德推脱的中介效应。依据以往标准，以下标准代表模型拟合良好(Hoyle & Panter, 1995)：χ^2/df 小于 5，RMSEA 和 SRMR 小于 0.08，TLI 和 CFI 大于 0.90。

4.1.5 结果与讨论

描述性统计和相关分析：A 参数、O 参数、道德推脱、身体攻击和口头攻击的描述统计和相关分析的结果如表 2 所示。具体来说，有中等证据支持 A 参数和 O 参数不相关(BF_{10} = 0.29)；A 参数和 O 参数分别与道德推脱、身体攻击呈显著负相关，而道德推脱、身体攻击、口头攻击两两之间呈显著正相关；口头攻击仅与 O 参数呈显著负相关，与 A 参数的相关不显著(r = -0.05, p = 0.349, BF_{10} = 0.09)。因此，在探究道德推脱在伤害行为/结果厌恶和攻击之间的中介作用时，不考虑道德推脱在 A 参数与口头攻击之间的中介效应。

表 2 描述统计及相关分析

变量	*M*	*SD*	1	2	3	4	5
1. A 参数	0.57	0.16	–				
2. O 参数	0.23	0.32	0.09	–			
3. 道德推脱	1.99	0.55	-0.17^{***}	-0.32^{***}	–		
4. 身体攻击	1.84	0.63	-0.16^{**}	-0.32^{***}	0.60^{***}	–	
5. 口头攻击	2.60	0.73	–0.05	-0.13^{**}	0.32^{***}	0.48^{***}	–

注：$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$。

中介效应分析：为揭示伤害厌恶抑制攻击的中介机制，本研究通过结构方程模型来检验(如图 2 所示)，在控制了性别和年龄后，道德推脱是否在 A 参数和身体攻击之间，以及在 O 参数和身体/口头攻击之间发挥中介作用。

中介模型的拟合良好，χ^2/df = 229.13/87, RMSEA [90% CI] = 0.06[0.05, 0.07], CFI = 0.94, TLI = 0.93, SRMR = 0.05。A 参数负向预测身体攻击的总效应显著(β = –0.12, p = 0.020, 95%CI = [–0.23, –0.02])，但对口头攻击(β = –0.04, p = 0.525, 95%CI = [–0.15, 0.07])的总效应不显著；O 参数负向预测身体攻击(β = –0.26, p <0.001, 95%CI = [–0.38, –0.12])和口头攻击(β = –0.14, p = 0.041, 95%CI =

[–0.27, –0.01])的总效应均显著。其中，A 参数、O 参数、道德推脱以及协变量预测身体攻击的 R^2 为 0.52，f^2 为 1.08，预测口头攻击的 R^2 为 0.14，f^2 为 0.16，达到了样本量可以检测到的最小效应量。这些结果支持了伤害厌恶可以抑制攻击的观点(Blair, 1995; Crockett et al., 2015; Yu et al., 2021)，与已有研究提示的伤害厌恶可以减少伤害行为的结果相符 (Contreras-Huerta et al., 2022; Crockett et al., 2010; Perera et al., 2016)。

实验 2a 采用 Bootstrap 法进行的间接效应检验发现，道德推脱在 A 参数与身体攻击(β = –0.08, p = 0.010, 95%CI = [–0.15, –0.02])之间发挥中介作用；道德推脱在 O 参数与身体攻击(β = –0.17, p < 0.001, 95%CI = [–0.27, –0.09])、口头攻击(β = –0.10, p = 0.002, 95%CI = [–0.18, –0.05])之间发挥中介作用；A 参数对身体攻击(β = –0.04, p = 0.343, 95%CI = [–0.11, 0.04])和口头攻击(β = 0.01, p = 0.821, 95%CI = [–0.09, 0.11])的直接效应不显著；O 参数对身体攻击(β = –0.09, p = 0.110, 95%CI = [–0.19, 0.02])和口头攻击(β = –0.04, p = 0.579, 95%CI = [–0.17, 0.09])的直接效应也不显著。上述结果与已有研究发现的道德情绪可以负向预测道德推脱(Chowdhury & Fernando, 2013; Ouvrein et al., 2018; Shen et al., 2023)以及道德推脱正向预测攻击(Dou et al., 2024; Gini et al., 2021; Teng et al., 2019)的结果一致。

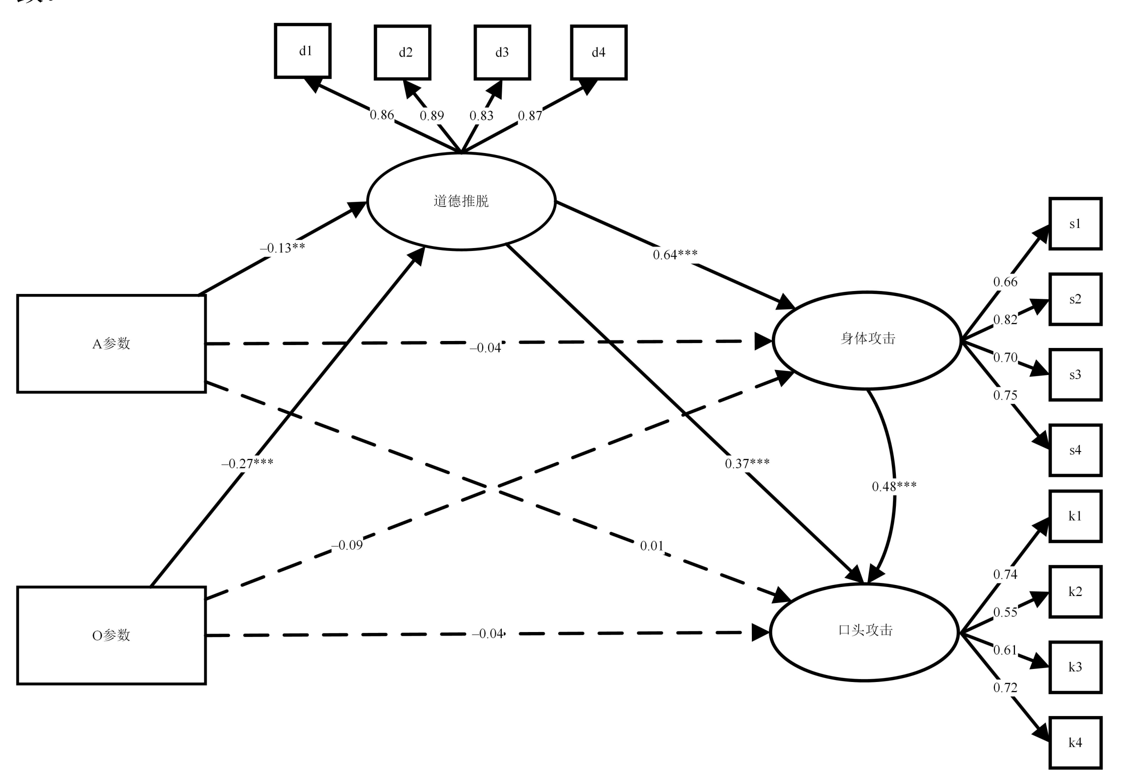


图 2 A/O 参数、道德推脱与身体/口头攻击的结构模型

注：路径系数为标准化系数，*p < 0.05; **p < 0.01; ***p < 0.001。

为了显示美观，性别和年龄的路径被隐藏，虚线表明路径不显著。

4.2 实验 2b 伤害行为/结果厌恶对攻击行为的抵御作用：道德推脱的中介作用

实验 2a 是采用问卷法来测量攻击，不能有效检验伤害行为/结果厌恶与攻击间的因果关系，因此实验 2b 采用实验任务来测量道德推脱和攻击行为。

4.2.1 被试

由于实验 2b 采用实验法测量道德推脱和攻击，因此参考 Cohen (1969)的建议，设置了中等效应量($f^2 = 0.15$)，预测变量为 5 个(即 A 参数、O 参数、道德推脱、性别、年龄)，求出达到 0.80 的统计效力需要的样本量为 92 人。本实验在正式实验前两周在本校招募了 124 名被试，并让其完成伤害行为/结果厌恶分离任务，根据分离工具的计算规则，排除了 11 人，以及在实验过程中不认真的被试 6 人，最终有效被试为 107 人(男生 33 名，女生 74 名)，年龄为 17.83~28.17 岁(M = 21.12, SD = 1.73)，均为右利手,视力或矫正视力正常，无精神病史，愿意接受电击实验，实验结束后根据被试选择和认真程度给予 20~35 元的报酬。随后，敏感性分析得到 107 人可检验的最小效应量为 $f^2 = 0.12$。

4.2.2 实验仪器与材料

实验程序采用 Eprime2.0 软件编写，在 Windows10 系统(分辨率 1280×1024)的电脑上完成。

伤害行为/结果厌恶分离任务：同实验 1 和实验 2a。

攻击-获得任务：该任务由疼痛-获得任务(pain-gain task, PGT；Feldmanhall et al., 2015)发展而来。在该任务中，被试(“决策者”)可以通过选择电击“接受者”来获得对应数量的金钱。此外，在电击选择之前增加了“道义恰当度”的评定任务，以测量被试对电击选择的道德推脱。该任务的具体实验流程如图 3 所示。

电刺激仪：实验 2b 使用商用电刺激仪(YRKJ-F1002)产生电刺激。通过两个直径为 1 厘米的 AgCl 电极可对被试施加不同强度的 2 秒时长的皮上电刺激。这两个电极之间的距离约为 2 厘米，固定在受试者的左手臂前侧上。在正式实验开始前实验人员用该仪器获得个体的疼痛阈限与感受，在实验之中和之后让被试相信其使用该电刺激仪电击了对方(实际并未发生)。

4.2.3 实验程序

被试到达实验室后的流程如下。第一步，被试阅读实验说明并填写知情同意书，研究者告知被试，实验中有“决策者”和“接受者”两个角色，决策者可以选择电击接受者，但接受者不能电击决策

者。被试的角色是决策者，可以通过选择电击接受者来获得额外的金钱。

第二步，我们使用电刺激仪和 9 级评定量表测量被试的疼痛阈限(“一接触就受不了的疼痛”)，并对所有被试进行疼痛校准。告知被试，接受者将接受带来同等主观疼痛程度的电击。

第三步，实验员让被试完成两轮练习实验，在每轮练习实验结束后均兑现电击和金钱。被试需要按电刺激仪的按键实施电击，但是不会真的通电。

第四步，被试完成正式实验。正式实验由 40 个正式试次和 10 个填充试次组成，分为两个部分，每个部分包含 25 个试次，采用随机顺序呈现。正式试次中电击时长的范围为 2~3 秒，金钱数额范围为 3~5 元，并将两种线索(电击时长与金钱数额)随机搭配。

正式实验的流程如图 3 所示：首先，屏幕中央呈现注视点“+”；之后，呈现由金钱数额和电击时长匹配的线索，被试思考本轮是否选择电击接受者，并评定在本轮选择电击接受者的道义恰当度(1 为“非常不恰当”，4 为“非常恰当”)；随后，再次呈现注视点“+”；接着，被试根据之前的思考按键决定是否选择电击，其中“⚡”代表电击，“▰”代表不电击，两个符号出现的位置(左右)随机；最后，反馈本轮获得的金钱数额。

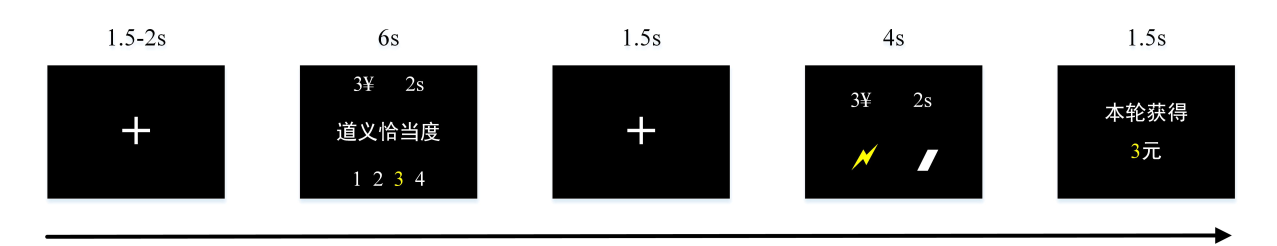


图 3 改编的疼痛-获得任务的单个试次示例

正式实验结束后，实验人员对被试进行访谈，并根据抽取到的试次等情况结算实验费。

4.2.4 统计处理

首先，我们借助 SPSS 24.0 对 A 参数、O 参数、道德推脱和攻击行为进行描述性统计和相关分析。对于 A 参数与 O 参数的关系以及相关不显著的变量关系，我们采用 JASP 进行贝叶斯相关分析。之后，我们借助 Process 插件采用 Bootstrap 5000 次重复抽样的方法检验道德推脱在 A/O 参数与攻击行为关系间的间接效应。

4.2.5 结果与讨论

描述性统计和相关分析：贝叶斯相关分析的结果显示，有较弱的证据支持 A 参数和 O 参数不相关($BF_{10} = 0.51$)，这与实验 1 和实验 2a 的结果一致。A 参数、O 参数、道德推脱、攻击行为的描述统计和相关分析结果如表 3 所示。第一，O 参数与道德推脱($r = –0.21, p = 0.028$)和攻击行为($r = –0.24, p = 0.014$)呈显著负相关，这与实验 2a 的结果一致。第二，A 参数与道德推脱($r = –0.02, p = 0.854, BF_{10} = 0.12$)、攻击行为 ($r = –0.01, p = 0.963, BF_{10} = 0.12$) 的相关均不显著，且贝叶斯因子提供了中等证据支持这一结果，这与实验 2a 的结果不一致。在接下来的中介效应分析中仅纳入 O 参数。

表 3 描述统计及相关分析

变量	*M*	*SD*	1	2	3	4
1. A 参数	0.51	0.16	–			
2. O 参数	0.25	0.27	0.17	–		
3. 道德推脱	2.00	0.58	–0.02	–0.21*	–	
4. 攻击行为	0.48	0.36	–0.01	–0.24*	0.60***	–

注：* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$。

中介效应分析：如图 4 和表 4 所示，O 参数可以显著负向预测道德推脱($\beta = –0.21, p = 0.008$)，而道德推脱可以正向预测攻击行为($\beta = 0.57, p < 0.001$)。Bootstrap 检验结果显示，O 参数对攻击行为的主效应显著($\beta = –0.23, p = 0.006$, 95%CI = [–0.39, –0.06])，但直接效应不显著($\beta = –0.12, p = 0.117$, 95%CI = [–0.26, 0.03])，道德推脱在 O 参数与攻击行为关系间的中介效应显著($\beta = –0.12, p = 0.013$, 95%CI = [–0.21, –0.02])，且占总效应的 50%，O 参数、道德推脱以及协变量预测攻击的 R^2 为 0.38, f^2 为 0.61，达到了样本量可以检测到的最小效应量。这些结果与实验 2a 的结果一致。

表 4 中介效应整体性检验

路径	效应量	95% CI		效应占比
		下限	上限	
总效应：O 参数→攻击行为	-0.23	–0.39	–0.06	100%
中介效应：O 参数→道德推脱→攻击行为	-0.12	–0.26	0.03	50.00%
直接效应：O 参数→攻击行为	-0.12	–0.21	–0.02	50.00%

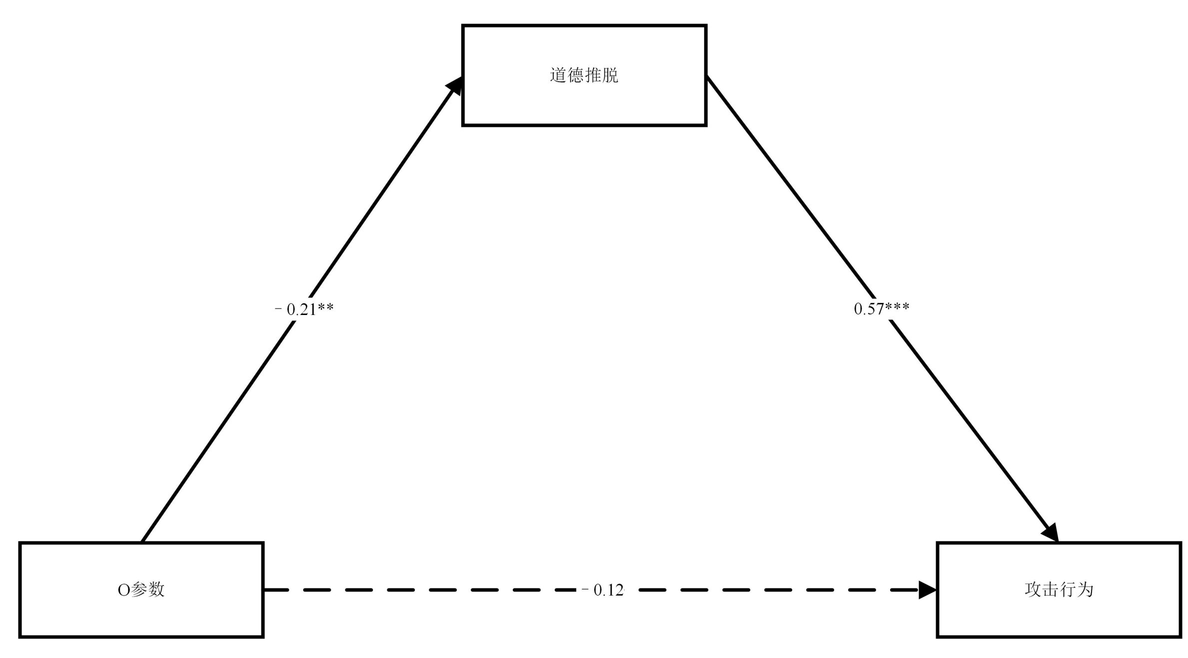


图 4 道德推脱在 O 参数与攻击行为之间的中介作用

注：* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$。

5 总讨论

为了探究伤害厌恶的结构及其抑制攻击的效应与路径，我们通过 1 个预实验开发了伤害行为/结果厌恶分离任务，并在此基础上实施了 3 个正式实验。具体发现如下：第一，A 参数(代表伤害行为厌恶)与 O 参数(代表伤害结果厌恶)的相关不显著。二者均与伤害行为厌恶问卷分数、共情、道德判断和道义主义参数正相关，与精神病态和身体攻击负相关。A 参数与伤害结果厌恶的问卷分数的相关不显著。第二，伤害结果厌恶可以通过道德推脱的中介作用负向预测攻击。第三，在实验 2a 中伤害行为厌恶可以通过道德推脱负向预测身体攻击，但在实验 2b 中没有发现这一关系。

5.1 伤害厌恶的结构

实验 1 的结果显示，代表伤害行为厌恶的 A 参数和代表伤害结果厌恶的 O 参数的相关较低而且不显著；A 参数与伤害结果厌恶的问卷分数的相关不显著；相比 A 参数，O 参数与共情关注、精神病态以及口头攻击的关系更强。这些结果支持了我们的假设，即伤害行为厌恶和伤害结果厌恶相互独立，且在性质与功能上存在明显差异(Ariso, 2022; Cushman et al., 2013; Miller et al., 2014)。

根据道德二元论(Gray et al., 2012)的观点，伤害行为厌恶和伤害结果厌恶的来源不同，这很可能导致二者的性质不同。例如，个体通常都会对伤害结果产生厌恶感，但对于没有出现或看到伤害结果的伤害行为就不一定有厌恶体验。因此，伤害结果厌恶可能更基本和普遍。例如，暴力抑制机制理论(Blair, 1995; Miller et al., 2014)指出，伤害结果厌恶是与生俱来的，更加根深蒂固。还有学者提出，伤害行为厌恶是在伤害结果厌恶的基础上通过强化学习形成的(Patil, 2015; Miller & Cushman, 2013)。伤害行为厌恶和伤害结果厌恶在性质上的这种差异很可能是导致它们的作用机制存在差异的根本原因。正因为伤害结果厌恶更基本和普遍，相对于伤害行为厌恶，其对攻击行为与心理的抑制作用更大、更稳健。

5.2 伤害行为/结果厌恶分离任务

本研究通过 PDP 范式对伤害行为厌恶和伤害结果厌恶进行分离，这一方法参考了以往使用行为指标测量心理现象的研究。例如，在认知心理学中，反应时和正确率等行为指标常被用来衡量认知变量(Liao et al., 2024; Pavailler et al., in press)，而在恐惧研究中，眨眼反应则被用于测量恐惧情绪(Khemka et al., 2017; Xia et al., 2024)。基于这一逻辑，本研究将对伤害行为和伤害结果的回避行为作为衡量伤害行为/结果厌恶情绪的行为指标，用以实现对这两类厌恶情绪的测量与计算。这一做法的合理性在于，伤害行为/结果厌恶的核心行为表现就是对伤害行为和结果的回避。

本研究采用 PDP 范式对伤害行为厌恶和伤害结果厌恶进行测量的优势主要有两个。首先，PDP 是采用行为指标来测量伤害行为/结果厌恶，可以避免被试自我报告产生的要求特征或社会期望效应，使得测量结果更客观和可靠。其次，使用该任务可以有效分离伤害行为厌恶和伤害结果厌恶，解决问卷法难以测量出相对单纯的伤害行为厌恶的问题。然而，该范式也存在一些缺陷。例如，该范式没有直接测量情绪体验，难以确定获得的 A 参数和 O 参数能否反映情绪体验。同时该任务无法检验题项间的同质性信度，因此我们后续的研究将采用重测信度指标对该任务的信度进行检验。此外，本研究发现 A 参数、O 参数与攻击等效标的相关不高，这可能由于影响攻击等效标变量的因素很多，伤害厌恶只是其中之一。

5.3 伤害结果厌恶抵御攻击的道德认知路径

实验 2a 和 2b 均发现道德推脱在伤害结果厌恶与攻击之间发挥中介作用，支持了研究假设。

5.3.1 伤害结果厌恶对攻击的抵御作用

实验 1 和 2(包括 2a 和 2b)的结果均显示，伤害结果厌恶与身体攻击、口头攻击均呈显著负相关；此外，伤害结果厌恶还可以抑制攻击行为。这些结果与已有研究中的观点以及获得的间接证据基本一致。例如，许多学者都认为伤害厌恶可以抑制攻击这一伤害行为(Blair, 1995, 2007; Cushman, 2013; Martinez et al., 2024)。此外，已有研究提供了一些间接证据，表明伤害厌恶可以抑制攻击行为。例如，听到受害者痛苦的声音就会减少电击行为(Buss, 1966)，5-羟色胺能促进伤害厌恶并抑制个体伤害其它玩家的行为(Crockett et al., 2010)。与以往研究相比，实验 2 提供了明确而直接的实证证据。这一结果的主要原因可能在于，攻击行为容易导致消极后果，而对伤害他人导致的消极后果的厌恶会引发对个体攻击行为的道德抑制反应，以避免这些消极后果的发生。

5.3.2 道德推脱的中介作用

实验 2 的两个子研究均发现，伤害结果厌恶可以负向预测道德推脱。这与已有的关于道德情绪与道德推脱关系的研究结果一致。例如，共情、内疚等道德情绪都被发现对道德推脱有抑制作用(Leviston & Walker, 2020; Ouvrein et al., 2018; Shen et al., 2023)。这一结果的主要原因可能在于，攻击行为通常都会导致伤害后果，而对伤害结果的厌恶所引发的道德抑制使得个体较难从道德上对攻击行为进行积极扭曲或美化。

本研究显示，道德推脱与攻击之间存在正向关系，这与大量研究获得的道德推脱会促进各种类型的攻击(Dou et al., 2024; Gini et al., 2021; Nocera et al., 2022)的结果相符。该结果出现的原因是，虽然道德系统会将攻击视为是错误的，从而抑制攻击(Bandura, 2002; Fitouchi et al., 2022; Hou et al., 2023)，但道德推脱却可以通过将攻击合理化而使道德系统的这种抑制减弱或失效(Bjärehed et al., 2019; Wang et al., 2022)。因此，道德推脱可以起到促进攻击的作用(Guo et al., 2024)。

道德推脱在伤害结果厌恶与攻击之间发挥中介作用的结果与已有研究发现的道德推脱在道德情绪(如共情、内疚)与攻击的关系中起中介作用的结果一致(Ouvrein et al., 2018; Wang et al., 2017; Shen et al., 2023)。这些结果提示，道德推脱是伤害厌恶抑制攻击的一个重要的道德认知机制。具体而言，伤害厌恶通过抑制个体对导致伤害后果的行为(如攻击)进行道德美化，从而降低了个体实施攻击的可能性或概率。

5.4 伤害行为厌恶与攻击、道德推脱的关系

5.4.1 伤害行为厌恶与攻击的关系

实验 1 和实验 2a 发现，伤害行为厌恶与身体攻击呈负相关关系，但与口头攻击的相关不显著；实验 2b 发现，伤害行为厌恶对攻击行为的抑制作用不显著。这些结果提示，伤害行为厌恶与攻击的关系似乎并不稳健。导致这一现象的原因可能有两个。第一，伤害行为厌恶(A 参数)强调的是在不考虑伤害结果的前提下，个体对伤害行为本身的厌恶。由于排除了伤害结果，这种道德情绪所引发的道德系统对攻击的道德抑制可能会比较弱，因此对攻击的抑制作用也较弱。第二，伤害行为的属性很可能会影响伤害行为厌恶诱发的道德抑制的强度。例如，有研究对比了人们在电车困境和人行桥困境中的反应，结果发现，在人行桥困境中只有少部分人会选择将一个人推下去来拯救更多人，而在电车困境中会有较多的人选择扳动开关，用一个人的性命换取更多人的性命(Cushman et al., 2006; Gawronski & Ng, 2025; Greene et al., 2009)。“推”是典型、直接和强烈的伤害行为，而“扳开关”则是不典型的伤害行为(Cushman et al., 2010; Miller et al., 2014)。因此，在人行桥困境中，个体的伤害行为厌恶可能引发了更强的对伤害行为的道德抑制，从而抑制了个体做出伤害行为选择。与身体攻击相比，口头攻击是一种不太典型的伤害行为，据此可以推测，伤害行为厌恶引发的对口头攻击的道德抑制较弱，因此伤害行为厌恶与口头攻击的相关可能不显著。类似地，与直接的身体攻击相比，通过按键做出电击选择属于不典型的间接性的身体攻击，而且在实验 2b 中仅抽取了部分试次来兑现，这可能是导致伤害行为厌恶与电击选择的相关不显著的主要原因。以上推测是否正确，还需要进一步的研究检验。

5.4.2 伤害行为厌恶与道德推脱的关系

实验 2a 的结果显示，伤害行为厌恶与道德推脱存在显著的负相关关系，但在实验 2b 中，伤害厌恶与道德推脱的相关则不显著。这提示，伤害行为厌恶与道德推脱的关系似乎并不稳健。这一现象的原因可能有以下两个。第一，伤害行为厌恶本身很难抑制个体对伤害行为进行道德认知上的积极扭曲。例如，当没有实际的伤害结果时，个体通常不会将伤害行为视为非常不道德的行为，因此更容易进行道德辩解。第二，伤害行为厌恶与道德推脱的关系还可能会受到研究方法的影响。实验 2a 采用了问卷法测量个体的道德推脱，而实验 2b 采用的是实验法测量个体在攻击任务中的道德推脱。实验 2b 中的按键选择属于不典型的伤害行为，同时被试不会直接看到任何伤害结果。在这种情况下，伤害行为厌恶对合理化电击选择行为的道德抑制作用可能就会相对较弱。

5.5 伤害行为/结果厌恶与不道德的认知和行为

综上所述，本研究发现，伤害行为/结果厌恶对道德推脱和攻击都有一定的抑制作用。这是因为，

根据道德系统会抑制不道德的心理和行为的观点(Bandura, 1999, 2002; Sonnentag et al., 2024)，作为道德系统的核心因素和主要的道德情绪，伤害行为/结果厌恶会对不道德的认知(如道德推脱)和行为(如攻击)产生抑制作用。

与此同时，伤害结果厌恶对道德推脱和攻击的抑制作用更稳健和明显，但伤害行为/结果厌恶的这种抑制作用则不太稳定，这提示伤害结果厌恶对不道德的认知和行为的抑制作用可能更大。出现这种结果的原因可能是，伤害结果厌恶更基本和普遍，因此其对道德推脱和攻击具有较稳健的抑制作用。

5.4 本研究的贡献与局限

本研究具有一定的理论意义和实践价值。首先，本研究采用 PDP 检验了伤害厌恶的二因素结构，发现伤害行为厌恶与伤害结果厌恶在抑制攻击认知与行为的效应上存在差异，并揭示了道德推脱在其中的中介作用。这些研究结果加深了我们对伤害厌恶的结构与功能的认识，对于攻击的预防与干预工作也有一定启发。例如，同时提升个体的伤害行为和伤害结果厌恶可能更有利于抑制个体的攻击认知和行为。其次，本研究使用加工分离程序开发的伤害行为/结果厌恶分离任务为伤害厌恶研究提供了一个有效的工具。

本研究也存在以下缺陷和问题。首先，本研究的伤害行为/结果厌恶分离任务主要涉及身体攻击的情境，而涉及口头攻击、间接攻击等情境较少，可能会导致低估伤害行为/结果厌恶与口头攻击、间接攻击的关系。未来研究可以开发涉及多种攻击种类的情境任务，并进一步检验伤害行为/结果厌恶在抑制不同种类的攻击上的异同。其次，如上所述，在本研究发现的伤害行为厌恶与道德推脱、攻击的关系存在矛盾的结果，未来研究有必要进一步探究这些变量之间的关系，尤其是要考察伤害行为厌恶预测或影响道德推脱、攻击的边界条件。

参考文献

Almeida, G., Flanagan, B., & Hannikainen, I. (2024). Trait empathy predicts purposivist rule application: Nationally representative survey evidence. *SSRN Electronic Journal*. Advance online publication. http://dx.doi.org/10.2139/ssrn.4710786

Ariso, J. M. (2022). Is there an internal link between seeing a human and seeing one to whom moral consideration is due? In C. Eriksen, J. Hermann, N. O'Hara, & N. Pleasants (Eds.), *Philosophical Perspectives on Moral Certainty* (pp.

212−228). Routledge.

Armstrong, J., Friesdorf, R., & Conway, P. (2018). Clarifying gender differences in moral dilemma judgments: The complementary roles of harm aversion and action aversion. *Social Psychological and Personality Science*, *10*(3), 353−363.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193−209.

Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, *31*(2), 101−119.

Bjärehed, M., Thornberg, R., Wänström, L., & Gini, G. (2019). Mechanisms of moral disengagement and their associations with indirect bullying, direct bullying, and pro-aggressive bystander behavior. *The Journal of Early Adolescence*, *40*(1), 28−55.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*, *57*(1), 1−29.

Blair, R. J. R. (2007). The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*, *11*(9), 387−392.

Buss, A. H. (1966). Instrumentality of aggression, feedback, and frustration as determinants of physical aggression. *Journal of Personality and Social Psychology*, *3*(2), 153−162.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63*(3), 452−459.

Caprara, G. V., Fida, R., Vecchione, M., Tramontano, C., & Barbaranelli, C. (2009). Assessing civic moral disengagement: dimensionality and construct validity. *Personality and Individual Differences*, *47*(5), 504−509.

Carroll, A., McCarthy, M., Houghton, S., Sanders O'Connor, E., & Zadow, C. (2018). Reactive and proactive aggression as meaningful distinctions at the variable and person level in primary school-aged children. *Aggressive Behavior*, *44*(5), 431−441.

Cen, Y., Su, S., Dong, Y., & Xia, L. X. (2022). Longitudinal effect of self-control on reactive-proactive aggression: mediating roles of hostile rumination and moral disengagement. *Aggressive Behavior*, *48*(6), 583−594.

Chowdhury, R., & Fernando, M. (2013). The relationships of empathy, moral identity and cynicism with consumers' ethical beliefs: the mediating role of moral disengagement. *Journal of Business Ethics*, *124*, 677−694.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press

Contreras-Huerta, L. S., Lockwood, P. L., Bird, G., Apps, M., & Crockett, M. J. (2022). Prosocial behavior is associated with transdiagnostic markers of affective sensitivity in multiple domains. *Emotion*, *22*(5), 820−835.

Conway, P., Dawtry, R. J., Lam, J., & Gheorghiu, A. I. (2024). Is it fair to kill one to save five? How just world beliefs shape sacrificial moral decision-making. *Personality and Social Psychology Bulletin*, Advance online publication. http://doi.org/10.1177/01461672241287815

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, *104*, 216−235.

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development*, *67*(3), 993−1002.

Crockett, M. J., Clark, L., Hauser, M. D., & Robbins, T. W. (2010). Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(40), 17433−17438.

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320−17325.

Crockett, M. J., Siegel, J. Z., Kurthnelson, Z., Ousdal, O. T., Story, G. W., Frieband, C., ... Dolan, R. J. (2015). Dissociable effects of serotonin and dopamine on the valuation of harm in moral decision making. *Current Biology*, *25*(14), 1852−1859.

Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review*, *17*(3), 273−292.

Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, *12*(1), 2−7.

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*(1), 6−21.

Cushman, F., Young, L., & Greene, J. D. (2010). Our multi-system moral psychology: towards a consensus view. In J. M. Doris (Ed.), *Oxford Handbook of Moral Psychology* (pp. 1−20). Oxford University Press.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychological Science*, *17*(12), 1082−1089.

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of*

Personality and Social Psychology, *44*(1), 113−126.

Deng, Q., Deng, J., Lai, H., Huang, Y., & Wang, M. (2017). The factor structure and psychometric properties of the Levenson self-report psychopathic scale. *Chinese Journal of Clinical Psychology*, *25*(04), 659−666.

[邓俏文, 邓嘉欣, 赖红玉, 黄语嫣, 王孟成. (2017). 莱文森精神病态自评量表的因子结构及心理测量学特性分析. *中国临床心理学杂志*, *25*(04), 659−661.]

Diedenhofen, B., & Musch, J. (2015). Cocor: a comprehensive solution for the statistical comparison of correlations. *Plos One*, *10*(4), Article e121945. http://doi.org/10.1371/journal.pone.0121945

Dou, K., Chen, L. Y., Li, Y. Y., Lin, X. Q., Yuan, X. Q., & Li, J. B. (2024). The longitudinal associations between nonviolent/violent stressful life events, moral disengagement, and online aggression in chinese young people through the lens of general aggression model. *Psychology of Violence*, *14*(5), 364−372.

Falla, D., Ortega-Ruiz, R., & Romera, E. M. (2023). Minimizing responsibility in the aggressive dynamics of bullying and its impact on other strategies of moral disengagement: A longitudinal study. *Current Psychology*, *42*(36), 32512−32523.

Feldmanhall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *Neuroimage*, *105*, 347−356.

Fitouchi, L., Andre, J. B., & Baumard, N. (2022). Moral disciplining: the cognitive and evolutionary foundations of puritanical morality. *Behavioral and Brain Sciences*, *46*, Article e293. http://doi.org/10.1017/S0140525X22002047

Fung, A. L. (2021). Sex differences in the relationships between forms of peer victimization and reactive and proactive aggression in schoolchildren. *International Journal of Environmental Research and Public Health*, *18*(10), Article 5443 http://doi.org/10.3390/ijerph18105443

Gawronski, B., & Ng, N. L. (2025). Beyond trolleyology: the cni model of moral-dilemma responses. *Personality and Social Psychology Review*, *29*(1), 32−80.

Gibson, S., & Ninness, B. (2005). Robust maximum-likelihood estimation of multivariable dynamic systems. *Automatica*, *41*(10), 1667−1682.

Gini, G., Thornberg, R., Bussey, K., Angelini, F., & Pozzoli, T. (2021). Longitudinal links of individual and collective morality with adolescents' peer aggression. *Journal of Youth and Adolescence*, *51*(3), 1−16.

Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, *23*(2), 206−215.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons:

The interaction between personal force and intention in moral judgment. *Cognition*, *111*(3), 364−371.

Guo, Y., Li, R., & Xia, L. X. (2024). Effects of relative deprivation on change in displaced aggression and the underlying motivation mechanism: A three-wave cross-lagged analysis. *British Journal of Psychology*, *115*(1), 1−19.

Hou, Y., Li, X., & Xia, L. X. (2023). Common mechanisms underlying the effect of angry rumination on reactive and proactive aggression: A moderated mediation model. *Journal of Interpersonal Violence*, *39*(5−6), 1035−1057.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications* (pp. 158−176). Thousand Oaks: Sage publications.

Jacoby, L. L. (1991). A process dissociation framework: separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513−541.

Khemka, S., Tzovara, A., Gerster, S., Quednow, B. B., & Bach, D. R. (2017). Modeling startle eyeblink electromyogram to assess fear learning. *Psychophysiology*, *54*(2), 204−214.

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151−158.

Leviston, Z., & Walker, I. (2020). Influence of moral disengagement on responses to climate change. *Asian Journal of Social Psychology*, *24*(2), 144−155.

Li, R., & Xia, L. (2021). The mediating effect of aggression motivation on the relationship between trait anger and reactive aggression: A longitudinal study. *Acta Psychologica Sinica*, *53*(7), 788−797.

[李芮，夏凌翔. (2021). 攻击动机对特质愤怒与反应性攻击关系的中介作用：一项纵向研究. *心理学报*, *53*(07), 788−797.]

Li, X., & Xia, L. (2024). A serial cascade effect of cybervictimization and hostile rumination on the within-person change of moral disengagement. *Journal of Personality. 92*(6), 1726−1743.

Li, Y. (2023). Cognitive and emotional factors in moral decision-making. *Journal of Education, Humanities and Social Sciences*, *22*, 774−779.

Liao, J., Wang, J., Jia, S., Cai, Z., & Liu, H. (2024). Correlation of muscle strength, working memory, and activities of daily living in older adults. *Frontiers in Aging Neuroscience*, *16*, Article1453527. http://doi.org/10.3389/fnagi.2024.1453527

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York, NY: Wiley.

Martinez, R. M., Chou, S. H., Fan, Y. T., Chen, Y. C., Goh, K. K., & Chen, C. (2024). Negative emotionality down-regulation affects moral choice but not moral judgement of harm: A pharmacological study. *Scientific Reports*,

14(1), Article 1200. https://doi.org/10.1038/s41598-024-51345-8

Miller, R. M., Hannikainen, I. A., & Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion*, *14*(3), 573−587.

Miller, R., & Cushman, F. (2013). Aversive for me, wrong for you: First-person behavioral aversions underlie the moral condemnation of harm. *Social and Personality Psychology Compass*, *7*(10), 707−718.

Nocera, T. R., Dahlen, E. R., Poor, A., Strowd, J., Dortch, A., & Van Overloop, E. C. (2022). Moral disengagement mechanisms predict cyber aggression among emerging adults. *Cyberpsychology-Journal of Psychosocial Research On Cyberspace*, *16*(1), 1−18.

Ogunfowora, B. T., Nguyen, V. Q., Steel, P., & Hwang, C. C. (2022). A meta-analytic investigation of the antecedents, theoretical correlates, and consequences of moral disengagement at work. *Journal of Applied Psychology*, *107*(5), 746−775.

Ouvrein, G., De Backer, C. J. S., & Vandebosch, H. (2018). Online celebrity aggression: A combination of low empathy and high moral disengagement? The relationship between empathy and moral disengagement and adolescents' online celebrity aggression. *Computers in Human Behavior*, *89*, 61−69.

Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Journal of Cognitive Psychology*, *27*(3), 349−366.

Pearson, K., & Filon, L. N. G. (1898). Mathematical contributions to the theory of evolution.—IV. On the probable errors of frequency constants and on the influence of random selection on variation and correlation. *Philosophical Transactions of the Royal Society of London. Series a, Containing Papers of a Mathematical Or Physical Character*, *191*, 229−311.

Perera, P., Canic, E., & Ludvig, E. A. (2016). Cruel to be kind but not cruel for cash: harm aversion in the dictator game. *Psychonomic Bulletin & Review*, *23*(3), 893−898.

Pavailler, N., Gevers, W., & Burle, B. (in press). Temporal metacognition: Direct readout or mental construct? The case of introspective reaction time. *Journal of Experimental Psychology: General*.

Quan, F., Yang, R., Zhu, W., Wang, Y., Gong, X., Chen, Y., ... Xia, L. (2019). The relationship between hostile attribution bias and aggression and the mediating effect of anger rumination. *Personality and Individual Differences*, *139*, 228−234.

Reynolds, C. J., & Conway, P. (2018). Not just bad actions: Affective concern for bad outcomes contributes to moral condemnation of harm in moral dilemmas. *Emotion*, *18*(7), 1009−1023.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331−363.

Rong, X., Sun, B., Huang, X., Cai, M., & Li, W. (2010). Reliabilities and validities of chinese version of interpersonal reactivity index. *Chinese Journal of Clinical Psychology*, *18*(2), 158−160.

[戎幸, 孙炳海, 黄小忠, 蔡旻颖, 李伟健. (2010). 人际反应指数量表的信度和效度研究. *中国临床心理学杂志*,*18*(2), 158−160.]

Sarlo, M., Lotto, L., Rumiati, R., & Palomba, D. (2014). If it makes you feel bad, don't do it! Egoistic rather than altruistic empathy modulates neural and behavioral responses in moral dilemmas. *Physiology and Behavior*, *130*, 127−134.

Shen, Y., Yuan, L., Xiong, X., & Xin, T. (2023). Empathy and cyberbystander behavior: The role of moral disengagement. *Current Psychology (New Brunswick, N.J.)*, *42*(21), 18070−18079.

Sonnentag, T. L., Wadian, T. W., & Wolfson, M. J. (2024). The role of moral identity in the salience of the prescriptive and proscriptive systems of moral self-regulation. *Ethics & Behavior*, *34*(6), 425−437.

Su, S., & Xia, L. X. (2024). Neurostructural correlates of harm action/outcome aversion: The role of empathy. *Neuroimage*, Advance online publication. https://doi.org/10.1016/j.neuroimage.2024.120972.

Teng, Z., Nie, Q., Guo, C., Zhang, Q., Liu, Y., & Bushman, B. J. (2019). A longitudinal study of link between exposure to violent video games and aggression in Chinese adolescents: The mediating role of moral disengagement. *Developmental Psychology*, *55*(1), 184−195.

Thornberg, R., Pozzoli, T., & Gini, G. (2022). Defending or remaining passive as a bystander of school bullying in sweden: The role of moral disengagement and anti-bullying class norms. *Journal of Interpersonal Violence*, *37*(19−20), NP18666−NP18689.

Vaughan, E. P., Speck, J. S., Frick, P. J., Walker, T. M., Robertson, E. L., Ray, J. V., ... Cauffman, E. (2023). Proactive and reactive aggression: developmental trajectories and longitudinal associations with callous-unemotional traits, impulsivity, and internalizing emotions. *Development and Psychopathology*, Advance online publication. http://doi.org/10.1017/S0954579423000317

Vollberg, M. C., & Cikara, M. (2024). Affective prediction errors in persistence and escalation of aggression. *Journal of Experimental Psychology-General*, *153*(6), 1551−1567.

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35−57.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426−432.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58−76.

Wang, X., Lei, L., Yang, J., Gao, L., & Zhao, F. (2017). Moral disengagement as mediator and moderator of the relation between empathy and aggression among Chinese male juvenile delinquents. *Child Psychiatry and Human Development*, *48*(2), 316−326.

Wang, X., Wang, W., Qiao, Y., Gao, L., Yang, J., & Wang, P. (2022). Parental phubbing and adolescents' cyberbullying perpetration: A moderated mediation model of moral disengagement and online disinhibition. *Journal of Interpersonal Violence*, *37*(7-8), NP5344−NP5366.

Wang, X., Yang, J., & Gao, L. (2013). The Chinese version of civic moral disengament scale. *Studies of Psychology and Behavior*, *11*(06), 838−842.

[王兴超, 杨继平, 高玲. (2013). 公民道德推脱问卷的中文版修订. *心理与行为研究*, *11*(06), 838−842.]

Wiech, K., Kahane, G., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2013). Cold or calculating? Reduced activity in the subgenual cingulate cortex reflects decreased emotional aversion to harming in counterintuitive utilitarian judgment. *Cognition*, *126*(3), 364−372.

Wu, Y., & Wen, Z. (2011a). The statistical analysis procedure involving null hypothesis significance testing. *Psychological Science*, *34*(1), 230−234.

[吴艳, 温忠麟. (2011a). 与零假设检验有关的统计分析流程. *心理科学*, *34*(1), 230−234.]

Wu, Y., & Wen, Z. (2011b). Item parceling strategies in structural equation modeling. *Advances in Psychological Science*, *19*(12), 1859−1867.

[吴艳, 温忠麟. (2011b). 结构方程建模中的题目打包策略. *心理科学进展*, *19*(12), 1859−1867.]

Xia, Y., Wehrli, J., Abivardi, A., Hostiuc, M., Kleim, B., & Bach, D. R. (2024). Attenuating human fear memory retention with minocycline: A randomized placebo-controlled trial. *Translational Psychiatry*, *14*(28), 1−8.

Yu, H., Siegel, J. Z., Clithero, J. A., & Crockett, M. J. (2021). How peer influence shapes value computation in moral decision-making. *Cognition*, *211*, Article 104641. https://doi.org/10.1016/j.cognition.2021.104641

Yu, H., Siegel, J. Z., & Crockett, M. J. (2019). Modeling morality in 3 - D: Decision - making, judgment, and inference.

Topics in Cognitive Science, *11*(2), 409−432.

Zapata, J., & Deroy, O. (2023). Ordinary citizens are more severe towards verbal than nonverbal hate-motivated incidents with identical consequences. *Scientific Reports*, *13*(1), Article 7126. https://doi.org/10.1038/s41598-023-33892-8

Zhao, H., & Seibert, S. E. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, *91*(2), 259−271.

The two-factor structure of harm aversion and the mechanism underlying its function of resisting aggression

CEN Yushan[1], XIA Lingxiang[1*], HUANG Runyu[2], LV Jie[1]

([1]*Research Center of Psychology and Social Development, Southwest University*, *Chongqing* 400715, *China*)

([2] *Department of Psychology, Faculty of Social Sciences, University of Macau*, *Macau* 519000, *China*)

Abstract

Harm aversion is a core component of human morality that prevents individuals from harming others and is crucial for biological evolution. It has the positive functions of reducing immoral behavior (such as aggression) and increasing prosocial behavior. Individuals who lack harm aversion may exhibit increased antisocial behavior, which in turn, may cause significant harm to other individuals and society in general. Therefore, exploring the concept of harm aversion has significant theoretical and practical value. However, it remains unclear whether harm aversion is composed of two relatively independent factors and how these two factors differ in their functions and mechanisms related to the inhibition of aggression. This study hypothesized that harm aversion comprises two independent factors: harm action aversion and harm outcome aversion. Additionally, both harm action aversion and harm outcome aversion can inhibit aggression, and moral disengagement is a crucial cognitive mechanism through which harm aversion inhibits aggression.

A pilot study and three formal experiments were conducted to test these hypotheses. In the pilot study (N=40), incongruent conditions and congruent conditions based on the process dissociation procedure were developed to separate harm action aversion and harm outcome aversion. Participants completed the harm action/outcome aversion dissociation task, assessed the difficulty of choice and imagination, and provided suggestions for modifications. In Experiment 1 (N=287), the modified harm action/outcome aversion dissociation task was used. The validity of this task was examined using criterion variables such as empathy, moral judgment, deontology parameter, psychopathy, physical aggression, and verbal aggression. Experiments 2a (N=423) and 2b (N=107) explored whether moral disengagement mediated the relationship between harm action/outcome aversion and aggression. Specifically, Experiment 2a used

questionnaires to measure moral disengagement and aggression, while Experiment 2b employed an aggression task to measure moral disengagement and aggressive behavior levels.

The pilot study demonstrated that the process dissociation procedure could effectively separate harm action aversion and harm outcome aversion. Experiment 1 found no significant correlation between harm action aversion (Parameter A) and harm outcome aversion (Parameter O). Both Parameters A and O were positively correlated with harm action aversion scores, empathy, moral judgment, and deontology parameter, and negatively correlated with psychopathy and physical aggression. Compared to Parameter A, Parameter O showed stronger relationships with empathy, psychopathy, and verbal aggression. The correlation between Parameter A and harm outcome aversion was not significant. In Experiments 2a and 2b, harm outcome aversion negatively predicted aggression through the mediating effect of moral disengagement. However, while Experiment 2a showed a mediating effect in that harm action aversion negatively predicted physical aggression through moral disengagement, this mediating effect was not significant in Experiment 2b.

These results suggest, therefore, that harm aversion comprises two independent factors: harm action aversion and harm outcome aversion. Harm outcome aversion can inhibit aggression through moral disengagement, but the inhibitory effect of harm action aversion on aggression and moral disengagement is not robust. In developing a research tool for harm aversion, testing the two-factor structure of harm aversion, and examining its cognitive and behavioral differences in inhibiting aggression, this study revealed that moral disengagement is an important cognitive pathway through which harm aversion inhibits aggression. The findings contribute to the theoretical development of harm aversion, and they advance research on the relationship between moral emotions and aggression. This has significant implications for the prevention and intervention of aggressive behavior, as well as practical applications in moral education.

Keywords harm action aversion, harm outcome aversion, process dissociation procedure, aggression, moral disengagement

附录

附录 1　预实验研究材料和程序

伤害行为/结果厌恶分离任务主要使用的是可以反映不相容和相容条件的情境材料。首先，我们基于 Miller 等人(2014)的伤害行为/结果厌恶问卷确定“仅表现伤害行为但不体现伤害结果”(如“用锤子砸一具尸体的手”、“按照电影剧本的情节，愤怒地咒骂一位老婆婆”、“用一把可伸缩的道具刀刺向一位演员的脖子”) 和“仅体现伤害结果但不表现伤害行为”(如“出现(吓人的)过敏反应”，“看见一个人在哭泣”)的典型例子和表述。其次，我们从道德两难情境(Armstrong et al., 2018; Conway & Gawronski, 2013)中挑选出可以兼备上述两种情况的故事情境(如“拷打”、“车祸情境”、“过敏情境”、“时光机情境”等)。再次，研究人员对 5 名非心理学专业的大学生进行访谈，要求其描述可能出现伤害行为和结果的情境(如“校园霸凌”、“恶意竞争”、“报复欺骗自己的人”)，由一名心理学教授与一名心理学专业的研究生对访谈结果进行讨论、分析和整理，并根据访谈材料补充编写兼备“仅表现伤害行为但不体现伤害结果”和“仅体现伤害结果但不表现伤害行为”这两种情况的故事情境。最后，我们对上述两个步骤编制出的故事情境进行梳理，合并同种类型的情境，最后获得了 32 个类型不同的故事情境。

随后预实验邀请 6 名心理学专业的研究生和本科生进行评定。首先，实验者邀请请他们对每个情境做出选择，然后评定每个情境的选择难易程度、想象难易程度，并对每个情境提出修改建议。在评定结果和修改建议的基础上，我们对这些情境材料进行多次讨论、筛选和修改。具体而言，依据 Conway 和 Gawronski (2013)研究中的选择比例，我们假设在不相容情境下选择伤害行为和伤害结果的比例大体相当，因此，我们对被试在选择中具有偏向性的选项的表述进行了修改，并进行了 3 次小样本(11~16 人)施测，从而使得选择不同的反应选项的比例基本相当，最终形成了 10 个情境材料，并分别命名被绑架作为人质、被命令拷打路人、选择穿越身份、完成特工任务、战场救人、表演剧本选择、帮朋友出气、报复骗子、卧底自保和应对竞争对手。

每个情境材料包括三个部分的内容：情境故事、不相容条件下的选项和相容条件下的选项。例如，“应对竞争对手”情境材料中的情境故事为“假如你在班上有一个竞争对手，他/她经常挑衅你，导致你心情很不好，你会做出什么选择？”该情境的不相容条件下的选项为“在被挑衅之后，你和朋友把发泄室的一个人偶沙包当作他/她来狠狠地打，边打边咒骂他/她”(注：该选项的性质为“实施不会直接看到伤害结果的伤害行为”；高伤害行为厌恶者则通常不做该选择，而高伤害结果厌恶者通常

会做该选择)和“你不出面，让朋友帮你出气，之后你连续好多天都看到他/她鼻青脸肿、精神萎靡地来上课”(注：该选项的性质为“看到伤害结果但不直接实施伤害行为”；高伤害行为厌恶者通常会做该选择，高伤害结果厌恶者则通常不做该选择)两个。该情境的相容条件下的选项为“和几个朋友找到他的社交账号，在评论区匿名辱骂他/她，之后的一周你都看到他因此情绪低落、萎靡不振的样子”(注：该选项的性质为“实施能看到伤害结果的伤害行为”；高伤害行为/结果厌恶者通常都不会做该选择)和“拿出自己所有的零花钱去辅导班学习，每天辛苦地学习 12 个小时以上，以此分散自己的注意力”(注：该选项的性质为“实施非伤害行为”；高伤害行为/结果厌恶者都会做该选择)。

预实验首先请被试在该 10 个情境故事的不相容和相容条件下分别进行选择，情境与条件采用伪随机的方式呈现(即相容条件或不相容条件不能连续出现超过两次，且同一种故事情境下的相容条件和不相容条件不能接连出现)，选项的顺序在被试间进行平衡，之后请被试评定在每个情境做出选择的难易程度、想象该情境的难易程度，并给出修改建议。选择难易程度和想象难易程度均为 7 点李克特量表计分，1 为“非常容易”，4，为“中等”，7 为“非常困难”。

附录 2　伤害行为/结果厌恶分离任务

序号	情境	相容条件	不相容条件
1	**人质困境**	1-1：你被一个变态的蒙面歹徒绑到一间木屋，你发现他还绑了其它人质，并对这些人质进行折磨和残害。歹徒告诉你，如果想要活着出去，你必须做出选择： 选择 1：用皮鞭把另一个被蒙住双眼的人质打到惨叫。 选择 2：你身上的所有积蓄和重要财物都被歹徒搜走。	1-2：你被一个变态的蒙面歹徒绑到密室，你发现他还绑了其它人质，并对这些人质实行不同的折磨和恐吓。歹徒告诉你，如果想要活着出去，你必须做出选择： 选择 1：用锤子在一个非常逼真的仿真假人身上钉 10 颗钉子，以恐吓其他人质。 选择 2：看着一个浑身是血的人质惨叫不止，持续 20 分钟以上。
2	**秘密行动**	2-1：请假想你是中东某秘密组织中的一名成员，这天你收到上级指示，要	2-2：请假想你是中东某秘密组织中的一名成员，这天你收到上级指示，要求

		求你严刑拷打一名路人，让他认罪并受 1 年的牢狱之灾。假定现在只有以下两种办法，你会选择： 选择 1：亲自用各种刑具拷打路人，使其痛苦到只能屈打成招。 选择 2：你违背上级指令，因此被关入禁闭室，在污秽的环境中，独自生活半个月。	你严刑拷打一名路人，让他认罪并受 1 年的牢狱之灾。假定现在只有两种办法，你会选择： 选择 1：你亲自动手，掌握好火候，持续低强度地拷打这名路人几天，再结合心理战术让这个路人在还没有真正受伤的时候就认罪。 选择 2：你让下属动手，在下属告诉你差不多的时候，你才走进刑房，看着这个被折磨得浑身是伤、痛苦得受不了的路人在你面前认罪。
3	**穿越身份**	3-1：假如你穿越到古代，并且只能从以下两个角色中选取一个，你会选择： 角色 A：你需要经常完成殴打他人的任务，并看到他人被你打到受伤。 角色 B：你每天需要没日没夜地干很多苦力活，例如打水、洗衣、烧柴、浇花等。	3-2：假如你穿越到古代，并且只能从以下两个角色中选取一个，你会选择： 角色 A：你需要经常在街上演出，演出内容通常是：用一把可伸缩的道具刀刺向同伴的脖子；用锤子砸向同伴的胸口等。 角色 B：你需要观察和记录受害人的受伤情况和死亡状态，并看管受害人的尸体。
4	**实习警察**	4-1：假如你是一名正在接受训练的特工，即将被派到恐怖分子的据点当卧底，为了检验你是否能胜任卧底任务，长官决定对你进行最后一次测试，他给了你两个选择，让你选择其中一个完成： 选择 1：亲自执行枪决任务，并看着犯人中枪躺在血泊中死去。	4-2：假如你是一名正在接受训练的特工，即将被派到恐怖分子的据点当卧底，为了检验你对敌人不会心慈手软，长官决定对你进行最后一次测试，他给了你两个选择，让你选择其中一个完成： 选择 1： 在同行穿着防弹衣的情况下，对不知情的同行开 5 枪(他/她不会真正

		选择 2：独自在缺乏食物来源和水源的荒野中生活一周。	受伤)。 选择 2：不动声色地观看 5 分钟的视频，视频呈现的是一个手筋脚筋被恐怖分子绞断的无辜者在痛苦呻吟、满地打滚。
5	**战场清理**	5-1：你是负责战场检查的士兵，你在清理战场时发现有一名敌军士兵还没有咽气，这时你会： 选择 1：补上一枪，看着他脖颈处流着鲜血，中弹身亡。 选择 2：放过这个敌军士兵，上级知道后，罚你背上全班的武器(大约 20 公斤)行军 10 公里。	5-2：你是负责清理战场的士兵，上级布置了两个战场清理任务，要求你必须选择一个去完成，你会选择： 选择 1：对可能潜藏有敌军士兵的山洞、草丛和战壕等地方进行机枪扫射或投掷手雷。 选择 2：在浑身是血、支离破碎的敌军士兵尸体上翻找有价值的情报或战利品。
6	**表演情境**	6-1：你是第一次参加试镜的演员，导演让你从下面两个剧本中选一个来试镜，你会选择： 剧本 1：你非常讨厌小 B，为了表现这种讨厌，这场戏需要你真的用力打小 B，小 B 的脸都会被你打肿。 剧本 2：你失恋了，非常痛苦，为了表现这种痛苦情绪，这场戏需要你浑身湿透的在雨中奔跑十分钟，然后扑倒在污泥中。	6-2：你是第一次参加试镜的演员，导演要求你扮演心狠手辣的人，为了顺利通过试镜，你会选择： 剧本 1：在 17 层高的楼顶上辱骂你的“竞争对手”，并毫不犹豫地把他从楼顶推下去(由于吊着威亚，对方不会受伤)。 剧本 2：在大牢中，面无表情地盯着一名四肢被绞断(化妆十分逼真)、不断哀嚎的“人质”，要他/她交出藏宝图。
7	**关系困境**	7-1：你的朋友被小 A 惹到了，十分生气，而你与小 A 没有交集，朋友想要做一些事来发泄自己的情绪，这时你会：	7-2：你的朋友被小 A 惹到了，十分生气，而你与小 A 没有交集，朋友想要做一些事来发泄自己的情绪，这时你会：

		选择 1：和朋友一起辱骂小 A，看着小 A 被骂哭。 选择 2：陪朋友去 KTV 唱歌发泄情绪，听他/她唱一晚上的歌，并因此错过你期待了 3 年的同学聚会。	选择 1：和朋友一起用针扎代表小 A 的布偶娃娃，并边扎边骂他/她。 选择 2：在朋友报复小 A 后，你多次看到小 A 躲在角落痛哭流涕的样子。
8	**江湖骗子**	8-1：假如你被一个谎称会算命的人拦了下来，他/她说你近期有灾，让你用 666 元买下一个护身符，而事后你发现自己被骗了，你会： 选择 1：你和其他被骗的人一起找到他/她，并辱骂、殴打他，让他/她被打的满脸是血、无地自容。 选择 2：通过兼职等努力工作的方式，把自己被骗的钱赚回来。	8-2：假如你被一个谎称会算命的人拦了下来，他/她说你近期有灾，让你用 666 元买下一个护身符，而事后你发现自己被骗了，你会： 选择 1：在他/她看不到你的地方，对着他/她不停地辱骂，诅咒他/她，将自己的怒火发泄出来。 选择 2：在听说他/她被其它受害人痛打了之后，你专门去看他/她被打断了腿、一瘸一拐的走在路上的样子，以消减怒火。
9	**卧底自保**	9-1：假如你是 A 组织的一名卧底，长期潜伏于 B 组织，并且有个下线负责保护你，一天你和下线悄悄正在 B 组织总部寻找机密资料，这时 B 组织的一名保镖似乎尾随着你们，并且离你们藏身的地方越来越近，为了不让他举报你，你会： 选择 1：你让下线弄出声响吸引保镖的注意力，你则从背后用小刀割开他的喉咙，使其咽喉流血倒地，不省人事； 选择 2：你拿出自己攒了一年的工资	9-2：假如你是 A 组织的一名卧底，长期潜伏于 B 组织，并且有个下线负责保护你，一天你和下线悄悄正在 B 组织总部寻找机密资料，这时 B 组织的一名保镖似乎尾随着你们，并且离你们藏身的地方越来越近，为了不让他发现你，你会： 选择 1：你让下线弄出声响吸引保镖的注意力，你则从背后用手掌劈向他的后脑勺，让他暂时昏迷。 选择 2：让下线找机会袭击这个保镖，自己悄悄离开。之后，你再和 B 组织

		来收买保镖，让他不要说出去。	的其他人一起看保镖躺在地上痛苦呻吟的样子。
10	**竞争对手**	10-1：假如你在班上有一个竞争对手，他/她经常挑衅你，导致你心情很不好，你会： 选择 1：和几个朋友找到他的社交账号，在评论区匿名辱骂他/她, 之后的一周你都看到他因此情绪低落、萎靡不振的样子。 选择 2：拿出自己所有的零花钱去辅导班学习，每天辛苦地学习 12 个小时以上，以此分散自己的注意力。	10-2：假如你在班上有一个竞争对手，他/她经常挑衅你，你很想出一口气，你会： 选择 1：在被挑衅之后，你和朋友把发泄室的一个人偶沙包当作他/她来狠狠地打，边打边咒骂他/她。 选择 2：你不出面，让朋友帮你出气，之后你连续好多天都看到他/她鼻青脸肿、精神萎靡地来上课。

附录 3 伤害行为/结果厌恶问卷

	完全没有	比较弱	有些弱	中等程度	有些强烈	比较强烈	非常强烈
1. 你拿着一把空枪对着你朋友的头，并扣动扳机(你和你的朋友都知道枪是空的)	1	2	3	4	5	6	7
2. 你按照电影剧本的情节，愤怒地咒骂一位老婆婆	1	2	3	4	5	6	7
3. 你用锤子砸一具尸体的手	1	2	3	4	5	6	7
4. 你按下静音键后，在电话里冲你妈妈大爆粗口	1	2	3	4	5	6	7
5. 你在自己最好朋友的背后，对着他(们)做一些下流的手势	1	2	3	4	5	6	7
6. 在一个演出中，你用一把可伸缩的道具刀刺向	1	2	3	4	5	6	7

一位演员的脖子							
7. 在一个逼真但安全的魔术表演中，你志愿上台“将一个女人锯成两半”	1	2	3	4	5	6	7
8. 在一个必需的医疗手术中，你切开一位病人的胃	1	2	3	4	5	6	7
9. 在朋友许可的情况下，你冲着防弹玻璃后的他(她)开了一枪	1	2	3	4	5	6	7
10. 你把一个看起来很逼真的婴儿娃娃砸在墙上	1	2	3	4	5	6	7
11. 你看见一个陌生人从楼梯上摔了下来	1	2	3	4	5	6	7
12. 你听见某人在没有使用麻醉的情况下拔牙	1	2	3	4	5	6	7
13. 你看到一位厨师不小心割破了手指	1	2	3	4	5	6	7
14. 你在晚间新闻中看到一个人溺水的镜头	1	2	3	4	5	6	7
15. 你看见一名足球运动员在比赛中摔断了腿	1	2	3	4	5	6	7
16. 你看见有人关轿车门时夹住了自己的手指	1	2	3	4	5	6	7
17. 你听到一个从自行车上摔下来受伤的男孩的哭喊声	1	2	3	4	5	6	7
18. 你看见一个油漆工从梯子上摔了下来	1	2	3	4	5	6	7
19. 你看见有人光着脚踩在玻璃碎片上	1	2	3	4	5	6	7
20. 你看见一个人不小心把开水洒到了自己的手上	1	2	3	4	5	6	7
21. 你看见一个穿凉鞋的女人不小心被购物车轧了自己的脚趾	1	2	3	4	5	6	7
22. 你看到一个男孩在吃了海鲜后出现(吓人的)过敏反应	1	2	3	4	5	6	7
23. 你听到一个被吓坏了的孩子在哭泣	1	2	3	4	5	6	7
24. 你看见一个女人做完一个很痛的手术后在哭泣	1	2	3	4	5	6	7

附录 4 道德两难任务

序号	场景	相容条件	不相容条件
1	时光机	你找到一台时光机，回到 1920 年。在入住酒店的时候，你遇到了一个年轻人。你知道这是乔治•布莱克曼，他后来绑架了一个孩子一周时间，直到孩子的家人付给他赎金。他住在你隔壁的酒店房间，门没有锁。在他睡觉的时候用枕头闷死他且不被人发现是很容易的，这样可以阻止他绑架孩子并索要赎金，但他现在还没有犯下任何罪行。 你为了阻止乔治绑架一名儿童而杀了他合适吗?	你找到一台时光机，回到 1920 年。在入住酒店时，你会遇到一位年轻士兵。你知道这是希特勒在德国纳粹党掌权前的样子。他住在你隔壁的酒店房间，门没有锁。在他睡觉的时候，用枕头把他闷死，然后不被人发现是很容易的，这样可以在二战开始和纳粹党建立之前就阻止他，但他还没有犯下任何罪行。 你为了防止第二次世界大战而杀死年轻的希特勒合适吗?
2	交通事故	当你开车穿过一条繁忙的城市街道时，一位抱着婴儿的年轻母亲突然走到你的行驶路线上。你开得太快以至于无法刹车，只能改变方向。不幸的是，唯一的转弯处现在有一群孩子。如果你为了避开年轻的母亲和婴儿而转弯，你会严重撞伤甚至撞死其中的一些孩子。 为了避开年轻母亲和婴儿而转弯撞向孩子合适吗?	当你开车穿过一条繁忙的城市街道时，突然一位抱着婴儿的年轻母亲走到你的行驶路线上。你开得太快以至于无法刹车，只能改变方向。不幸的是，你唯一能转弯的地方现在正坐着一位老太太。你为了避开年轻的母亲和婴儿而转弯，你会撞伤甚至撞死老太太。 为了避开年轻母亲和孩子，而转弯撞倒老太太合适吗?
3	贫困生活	你是一个贫困家庭的户主，你有一个 8 岁和一个 10 岁的儿子，你让他们每天在家里的农场里劳作，以	你是一个贫困家庭的户主，你的庄稼连续第二年歉收，你已经没有办法养活家人了。你有一个 8 岁和一

		此来教导他们努力工作的价值。然而，由于繁重的搬运任务，你的女儿不能在农场工作。你担心她不工作就不知道赚钱的不易。你认识一个男人，他来自你的村庄，住在城市里，他拍摄露骨的色情电影，主角是像你女儿这样的女孩。他当着你女儿的面告诉你，在他的工作室工作一年，你女儿就可以赚到很多钱，足够你们一家吃好几年。 为了让你的女儿明白赚钱的不易，你让她拍摄色情电影合适吗?	个 10 岁的儿子还太小，不能外面工作，但你的女儿可以去赚钱。你认识一个男人，他来自你的村庄，住在城市里，他拍摄露骨的色情电影，主角是像你女儿这样的女孩。他当着你女儿的面告诉你，在他的工作室工作一年，你女儿就可以赚到很多钱，足够你们一家吃好几年。你为了养家糊口而让你女儿拍摄色情电影合适吗?
4	**哭泣的婴儿**	现在是战争时期。敌军士兵占领了你们的村庄。他们接到命令，要俘虏所有剩下的平民，让他们在矿井里采石。你和一些村民躲在一个地窖里，外面有士兵的声音，他们来搜查房子里的贵重物品。一个孤儿开始大哭，你捂住她的嘴，不让她发出声音。如果你把手从她嘴里拿开，哭声会引起士兵们的注意，他们就会抓住你和其他人。为了让你和其它人不在矿井里劳作，你必须把孩子闷死。 为了不让你自己和其他村民被抓，你闷死孩子合适吗?	现在是战争时期。敌军士兵占领了你们的村庄。他们奉命杀死所有剩余的平民。你和一些村民躲在一个地窖里，外面有士兵的声音，他们来搜查房子里的贵重物品。一个婴儿开始大哭，你捂住她的嘴，不让她发出声音。如果你把手从婴儿的嘴里拿开，她的哭声会引起士兵的注意，他们会杀死你和其他人。为了救你和村民，你必须把孩子闷死。为了不让你和村民被杀，你闷死孩子合适吗?
5	**关系困境**	你现在离婚了，你的前妻(前夫)住在很远的地方，你很长时间都不会见	你现在离婚了，你的前妻(前夫)住在很远的地方，你很长时间都不会见

		到你的前妻(前夫)。然而， 你最近开始和一个非常棒的人约会，你的新伴侣沉浸在爱情中，你对他(她)的感情也很强烈。你想尽全力让这段关系顺利发展，并相信这将是第二次婚姻。不幸的是，你的前妻(前夫)传染给你一种非常严重且无法治愈的性传播疾病。你的新伴侣不知道这件事。你如果假装不知道这种性病，现在的关系就可以继续。 为了让她/他开心，让你们的关系持续下去，对你的新伴侣隐瞒你的性病合适吗?	到你的前妻(前夫)。然而， 你最近开始和一个非常棒的人约会，你的新伴侣沉浸在爱情中，你对他(她)的感情也很强烈。你想尽全力让这段关系顺利发展，并相信这将是第二次婚姻。不幸的是，你的新伴侣已经告诉你很多次了，他(她)不能接受和离婚的人约会，因为这在他(她)的宗教信仰中是禁止的。你的伴侣不知道你的前妻(前夫)。你如果假装自己从未离过婚，你现在的关系就可以继续下去。 为了让她/他开心，让你们的关系继续下去，对你的新伴侣隐瞒前一段感情合适吗?
6	**流产困境**	你是一名医生，一个年轻女孩怀孕了，但她还没有准备好要孩子。她没有读完高中，没有收入，被孩子父亲抛弃。如果她现在生了孩子，就只能靠国家福利生活了，这会让她和孩子的日子艰难。她认为以后再生孩子会更明智。尽管这对她来说很困难，她还是要求你把孩子打掉。 为了让女孩过上更好的生活，你进行人流合适吗?	你是一名医生，一个年轻女孩怀孕了，但是她的身体出现了异常的生理反应。她患上了严重的癫痫，这是一种危险的疾病，会导致血压迅速升高。唯一的治疗方法就是剖腹产。除非孩子尽快出生，否则母亲会死亡。然而，婴儿还太小，无法独立生存，如果它现在出生就会死亡。所以，尽管这对她来说很困难，这位女孩还是要求你打掉孩子。 为了挽救女孩的生命，你进行人流合适吗?
7	**拷打困境**	你是一名警官，最近抓住了一名罪	你是一名警官，最近抓住了一名罪

		犯。据说他操纵了一系列爆炸装置:一些已经爆炸，一些尚未引爆。他在城市咖啡馆外放置炸药，让它们在周围没人的时候爆炸。他的炸药装在油漆罐里，这样就能向附近的物体喷射油漆。他已经在许多咖啡馆喷射了油漆，但可能还有其它炸药。现在罪犯被拘留了，你想知道未爆炸的炸弹在哪里，这样你就能拆除它们。他拒绝说话，你可以使用“攻击的审讯技巧”，比如把他的头压在水里打他。 为了找到并拆除未爆炸的炸弹，你使用“攻击的审讯技巧”合适吗?	犯。据说他操纵了一系列爆炸装置:一些已经爆炸，一些尚未引爆。他在城市咖啡馆外放置了炸弹，并在人们在露台上喝咖啡的时候引爆。他已经伤害了很多人，可能还会伤害更多人。现在罪犯被拘留了，你想知道未爆炸的炸弹在哪里，这样你就能拆除它们。他拒绝说话，你可以使用“攻击的审讯技巧”，比如把他的头压在水里打他。 为了找到并拆除未爆炸的炸弹，你使用“攻击的审讯技巧”合适吗?
8	**疫苗政策**	你是一名医生，诊所里挤满了感冒的病人。你刚收到一批可以治愈流感的药物但这些药物有严重的副作用。如果你给你的病人服用这种药物，一小部分人会死于副作用，但大多数人会活下来。如果你不这样做，大多数人的感冒在一段时间内都无法痊愈。 你给你的病人用药合适吗?	你是一名医生，诊所里挤满了重病患者，你刚收到一批药可以治愈这种病，但这些药有严重的副作用。如果你给你的病人服用这种药物，一小部分人会死于副作用，但大多数人会活下来。如果你不这样做，大多数人会死于这种疾病。 你给你的病人用药合适吗?
9	**动物研究**	你受雇于一家制药公司，对他们的产品进行研究。由于产品必须适合人类使用，它们首先在动物身上 进行试验。你的工作是找出各种化学物质对老鼠、鸽子、兔子和猴子	你受雇于一家制药公司对他们的产品进行研究。由于产品必须适合人类使用，它们首先在动物身上进行试验。你的工作是找出各种化学物质对老鼠、鸽子、兔子和猴子的影

		的影响。有些化学物质会引起动物不适，甚至是永久性的伤害。你正在研究的化学物质将成为一种新的祛痘洗面奶的一部分，这将给油性皮肤的人带来新的治疗方案。你可以预期，使用这些化学物质会让很多人对自己的皮肤变好。 你在动物身上试验这些化学药品合适吗?	响。有些化学物质会引起动物不适，甚至是永久性的伤害。你正在研究的化学物质将成为一种新的艾滋病药物的成分，它将给世界各地数百万艾滋病患者带来新的希望。你想用化学药品拯救很多人的生命。 你在动物身上试验这些化学药品合适吗?
10	**边境困境**	你是一名边境守卫，在祖国和一个饱受战乱国家的边境上站岗。你注意到一个年轻人开着一辆廉价汽车向检查站走来，脸上带着坚定的表情。你怀疑他打算越过边境在你的国家进行非法工作。他正在迅速地接近你的车站。 你射杀正在靠近的人合适吗?	你是一名边境守卫，在祖国和一个饱受战乱国家的边境上站岗。你注意到一个年轻人开着一辆廉价汽车向检查站走来，脸上带着坚定的表情。你怀疑他想炸毁检查站，杀死里面所有的士兵。他正在迅速地接近你的车站。 你射杀正在靠近的人合适吗?

附录 5 共情分量表

	非常不符合	**比较不符合**	**不确定**	**比较符合**	**非常符合**
1. 对那些不幸的人，我通常会感到心疼和担忧	1	2	3	4	5
2. 我不会同情遇到麻烦的人(R)	1	2	3	4	5
3. 当看到有人被欺负时，我就想去保护他(们)	1	2	3	4	5

4. 他人的不幸通常*不会*让我心绪不安(R)	1	2	3	4	5
5. 当看到他人被不公正对待时，我并不会多么同情他(们)(R)	1	2	3	4	5
6. 我常常会触景生情	1	2	3	4	5
7. 我认为自己是一个相当心软的人	1	2	3	4	5

附录 6 莱文森精神病态量表

	完全不符合	不符合	符合	完全符合
1. 成功取决于适者生存 我不关心失败者	1	2	3	4
2. 我发现自己总是周而复始的陷入同一种麻烦中	1	2	3	4
3. 对我来说，所谓的正确就是能避开不好的结果	1	2	3	4
4. 我常常感到厌烦	1	2	3	4
5. 在今天这个世界，我觉得为了取得成功，做任何事都是正当的	1	2	3	4
6. 我人生的主要目标是尽可能多地得到好处	1	2	3	4
7. 赚大钱是我最重要的目标	1	2	3	4
8. 我很快就对开始做的事失去兴趣	1	2	3	4
9. 我让别人去担心崇高品格， 我只考虑什么对我有好处	1	2	3	4
10. 有些人很蠢，上当受骗了也是活该	1	2	3	4
11. 我曾经多次与别人大声争吵	1	2	3	4
12. 我会说人们想听的话，因为这样他们就会愿意为我做事	1	2	3	4

13. 我通常用发怒来发泄我的挫败感	1	2	3	4
14. 我经常会欣赏某个非常机智的骗局	1	2	3	4
15. 我强调，在追求自己的目标时尽可能不伤害到其他人(R)	1	2	3	4
16. 我喜欢操控他人的情感	1	2	3	4
17. 如果我的言行导致某人感到痛苦， 我心里会不好受(R)	1	2	3	4
18. 即使我很努力的想要销售某个东西，我也不会撒谎(R)	1	2	3	4
19. 作弊是不正当的，因为那样对他人不公平(R)	1	2	3	4

附录 7 身体攻击与口头攻击分量表

	非常不符合	比较不符合	不确定	比较符合	非常符合
1. 有时我没能控制住冲动就打了人	1	2	3	4	5
2. 我不同意朋友意见时，就当面反对	1	2	3	4	5
3. 如果受到多次挑衅，我就会打人	1	2	3	4	5
4. 我发现，自己常常与人争论	1	2	3	4	5
5. 我总能找到可以去打人的理由	1	2	3	4	5
6. 当有人来烦我的时候，我就容易骂人	1	2	3	4	5
7. 我比一般的人打架要多一些	1	2	3	4	5
8. 当有人与我意见不同时，我会忍不住与其争论	1	2	3	4	5

9. 如果确有必要，我会用打人的方法来维护自己的权利	1	2	3	4	5
10. 我的朋友们都认为我有点喜欢争论	1	2	3	4	5
11. 我会和仅仅挤了我一下的人打起来	1	2	3	4	5
12. 要是有人打我，我就打回去	1	2	3	4	5
13. 我会威胁我认识的人(要打他/她)	1	2	3	4	5
14. 我当着别人的面打砸东西	1	2	3	4	5

附录 8 公民道德推脱问卷

	完全不同意	比较不同意	不确定	比较同意	完全同意
1. 既然还没有完善的垃圾处理设施，那么责备把垃圾放到街上的人就没有道理	1	2	3	4	5
2. 一些人就不应该存活在世上	1	2	3	4	5
3. 忘了填写发票信息也没什么，因为检查这类失误本身就是财务人员的责任	1	2	3	4	5
4. 既然都有人做了更严重的破坏公物的行为，那么就没有理由去处罚那些仅在墙上涂鸦的人	1	2	3	4	5
5. 当路上的车都开得很快时，为了跟上其他车而超速的司机就不应该受到处罚	1	2	3	4	5
6. 既然是整个社会导致了环境恶化，那么个人对这一问题的担心就显得没有意义了	1	2	3	4	5
7. 考虑到国库的钱多半都被浪费掉了，那么避税行为就不应该受到谴责	1	2	3	4	5

8. 那些行为粗野的人也该被粗野地对待	1	2	3	4	5
9. 在大型百货商场偷点东西对其收入(盈利)的影响很小	1	2	3	4	5
10. 那些受害人通常没有掌握远离伤害的方法	1	2	3	4	5
11. 因为保险公司会对损失进行赔偿，所以失窃不会对零售商带来太大的损失	1	2	3	4	5
12. 在墙上涂鸦是在表现创新精神	1	2	3	4	5
13. 如果我们在一次伤害事件中仅起着次要作用，那么无需为此内疚	1	2	3	4	5
14. 经济交易中的欺诈只是一种战略	1	2	3	4	5
15. 用强硬措施让那些讨厌的人安静下来是合情合理的	1	2	3	4	5
16. 当所有人都违反规则时，就不应该责备个体	1	2	3	4	5
17. 赌博与其他活动一样，都是一种消遣	1	2	3	4	5
18. 为了科学进步，在高风险的实验中像“小白鼠”那样使用人类也是合法的	1	2	3	4	5
19. 如果一个人忘了拿自己的随身物品，那么这些东西被偷就是他自己的责任	1	2	3	4	5
20. 如果一个人在争吵中失控，那么他不需要为自己行为后果负全责	1	2	3	4	5
21. 相比工业造成的巨大污染，在街上乱扔垃圾不应受到惩罚	1	2	3	4	5
22. 为了保护自己的利益，通常必然会使用到武力	1	2	3	4	5
23. 既然社会上腐败盛行，那么任何人都推辞不掉别人给的好处	1	2	3	4	5
24. 为了维护家庭团结，即使犯了重罪的家人也应该被保护起来	1	2	3	4	5
25. 破坏老旧设施是说服政府提供新设施的一种方式	1	2	3	4	5
26. 既然是为了追求速度而制造汽车，那么超速就不是司机的错	1	2	3	4	5
27. 既然不少成人都使用麻醉品(吸毒)，那么青少年聚在一起抽烟就不该被认为是犯了多大的错	1	2	3	4	5
28. 对手(敌方)就该被羞辱和粗暴对待	1	2	3	4	5

29. 忠诚包括不告发犯了罪的朋友	1	2	3	4	5
30. 员工不该为执行了老板的违法指示而负责	1	2	3	4	5
31. 只有把某些人像牛马那样对待，才能驱动他们去工作	1	2	3	4	5
32. 色情读物也是普及性知识的一种方式	1	2	3	4	5